CENGAGE Learning

UNITEC
Universidad Tecnológica de México

Estadística inferencial aplicada

Mendenhall • Beaver • Beaver

Anderson • Sweeney • Williams • Camm • Cochran

UNITEC 50 AÑOS

Estadística inferencial aplicada

William Mendenhall
Robert J. Beaver
Barbara M. Beaver
David R. Anderson
Dennis J. Sweeney
Thomas A. Williams
Jeffrey D. Camm
James J. Cochran

Traducción

Jorge Alberto Velázquez Arellano
Lorena Peralta Rosales

Revisión técnica

M. I. Ángel Leonardo Bañuelos Saucedo
Coordinador General del Centro de Docencia
"Ing. Gilberto Borja Navarrete"
Profesor de Carrera Titular
Facultad de Ingeniería
Universidad Nacional Autónoma de México (UNAM)

María de Guadalupe Arroyo Santisteban
Academia de Matemáticas ECEE
Universidad Panamericana

Ignacio García Juárez
Academia de Matemáticas ECEE
Universidad Panamericana

José Cruz Ramos Báez
Academia de Matemáticas ECEE
Universidad Panamericana

Iren Castillo Saldaña
Academia de Matemáticas ECEE
Universidad Panamericana

Revisión técnica para esta adaptación

Omar González González
Director Académico de Matemáticas e
Ingeniería Industrial y de Sistemas
UNITEC, campus Cuitláhuac

CENGAGE Learning®

Australia • Brasil • Corea • España • Estados Unidos • Japón • México • Reino Unido • Singapur

CENGAGE Learning®

Estadística inferencial aplicada
William Mendenhall, Robert J. Beaver, Barbara M. Beaver, David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran

Gerente Editorial de Contenidos en Español:
Jesús Mares Chacón

Editora de Adquisiciones para Latinoamérica:
Claudia C. Garay Castro

Gerente de Manufactura para Latinoamérica:
Antonio Mateos Martínez

Gerente de Desarrollo Editorial en Español:
Pilar Hernández Santamarina

Gerente de Proyectos Especiales:
Luciana Rabuffetti

Coordinador de Manufactura:
Rafael Pérez González

Editora:
Cinthia Chávez Ceballos

Diseño de portada e interiores:
Ediciones OVA

Imagen de portada:
© Katty2016/Shutterstock

Fotografías de contraportada:
© UNITEC

Composición tipográfica:
Ediciones OVA

© D.R. 2018 por Cengage Learning Editores, S.A. de C.V., una Compañía de Cengage Learning, Inc.
Carretera México-Toluca 5420, Oficina 2301.
Colonia El Yaqui, Delegación Cuajimalpa,
Ciudad de México, C.P. 05320
Cengage Learning® es una marca registrada usada bajo permiso.

DERECHOS RESERVADOS. Ninguna parte de este trabajo amparado por la Ley Federal del Derecho de Autor, podrá ser reproducida, transmitida, almacenada o utilizada en cualquier forma o por cualquier medio, ya sea gráfico, electrónico o mecánico, incluyendo, pero sin limitarse a lo siguiente: fotocopiado, reproducción, escaneo, digitalización, grabación en audio, distribución en Internet, distribución en redes de información o almacenamiento y recopilación en sistemas de información, a excepción de lo permitido en el Capítulo III, Artículo 27h de la Ley Federal del Derecho de Autor, sin el consentimiento por escrito de la Editorial.

Esta es una adaptación de los libros:
Introducción a la probabilidad y estadística, 14a. edición, de William Mendenhall, Robert J. Beaver y Barbara M. Beaver, publicado por Cengage Learning con ISBN 978-607-519-876-7, traducido de la obra *Introduction to Probability and Statistics*, 14th Edition, publicado en inglés por Cengage Learning © 2013, ISBN 978-113-310-375-2, y *Estadística para negocios y economía*, de David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm y James J. Cochran, publicado por Cengage Learning con ISBN 978-607-522-515-9, traducido de la obra *Statistics for Business & Economics*, 12th Edition, publicado en inglés por South-Western, Cengage Learning © 2014, ISBN: 978-113-327-453-7

Datos para catalogación bibliográfica:
Mendenhall, William, Robert J. Beaver, Barbara M. Beaver, David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, James J. Cochran.
Estadística inferencial aplicada
ISBN: 978-607-526-560-5

Visite nuestro sitio en:
http://latinoamerica.cengage.com

Impreso en México
1 2 3 4 5 6 20 19 18 17

Contenido

1 **TEORÍA DEL MUESTREO**

1.1 Concepto de muestra 2
1.2 Características de la muestra 2
1.3 Tipos de muestreo 2
Muestreo aleatorio 2
Muestreo estratégico 4
Muestreo sistemático 5
Muestreo de conglomerados 5
Ejercicios 6

2 **DISTRIBUCIONES MUESTRALES**

2.1 Conceptos fundamentales 9
2.2 Distribución muestral de la media, propiedades, error estándar de la media 11
2.3 Error estándar 11
2.4 Teorema del límite central 14
2.5 Distribución muestral de la proporción 17
Ejercicios 21

3 **INFERENCIA ESTADÍSTICA CON MUESTRAS GRANDES**

3.1 Inferencia estadística, tipos de estimadores, evaluación de la bondad de un estimador puntual y de un estimador por intervalo 24
Inferencia estadística 24
Tipos de estimadores 25
3.2 Estimación puntual de la media de una población, estimación por intervalo de la media de una población 26
Estimación por intervalo 31
3.3 Estimación de la diferencia de dos medias 38
3.4 Estimación del parámetro de una población binomial, estimación de la diferencia entre dos proporciones 42
3.5 Selección del tamaño de la muestra 45
3.6 Utilización de software de estadística para la solución de casos 48
Ejercicios 51

4 **INFERENCIA ESTADÍSTICA CON MUESTRAS PEQUEÑAS**

4.1 La distribución *t* de Student 54
Suposiciones detrás de la distribución *t* de Student 57

4.2 Inferencia con muestras pequeñas para la media de la población 57
4.3 Inferencia con muestras pequeñas para la diferencia entre dos medias 62
4.4 Prueba para diferencias pareadas 68
4.5 Inferencias acerca de la varianza de una población 73
4.6 Comparación de las varianzas de dos poblaciones 79
Prueba y estimación de muestra pequeña: Microsoft Excel 81
4.7 Utilización de software de estadística para la solución de casos 81
Prueba y estimación de muestra pequeña: *MINITAB* 83
Ejercicios 87

5 PRUEBA DE HIPÓTESIS
5.1 Definición, prueba de hipótesis de uno y dos extremos 90
5.2 Tipos de hipótesis: nula y alternativa 90
5.3 Nivel de significancia, errores tipo I y tipo II 91
5.4 Prueba de hipótesis para medias y proporciones 92
Prueba de hipótesis e intervalos de confianza 94
Prueba para una proporción binomial 95
Prueba para la diferencia entre dos proporciones binomiales 98
5.5 Prueba de hipótesis para una y dos varianzas 101
5.6 Utilización de software de estadística para la solución de casos 105
Ejercicios 108

6 ANÁLISIS DE VARIANZA
6.1 Definición, conceptos fundamentales 111
¿Qué es un análisis de varianza? 112
Suposiciones para un análisis de varianza 112
6.2 Análisis de varianza de una vía o factor 113
Prueba de hipótesis 114
Variación entre grupos y dentro de los grupos 117
La distribución *F* 119
Estadística de prueba, regla de decisión 121
6.3 Análisis de varianza de dos vías 123
Pruebas de hipótesis 124
Variaciones considerando grupos y bloques 124
Estadísticos de prueba, regla de decisión 127
6.4 Utilización de software de estadística para la solución de casos 130
Procedimientos de análisis de varianza: *Microsoft Excel* 130
Ejercicios 131

7 PRUEBAS DE CHI CUADRADA
7.1 Pruebas de chi cuadrada para la diferencia entre dos proporciones (muestras independientes) 135
7.2 Pruebas de chi cuadrada para las diferencias entre más de dos proporciones 136
7.3 Prueba de independencia con chi cuadrada 138

7.4 Utilización de software de estadística para la solución de casos 142
La prueba chi cuadrada: *Microsoft Excel* 142
La prueba chi cuadrada: *MINITAB* 143
Ejercicios 146

8 REGRESIÓN LINEAL Y CORRELACIÓN

8.1 Diagrama de dispersión, el modelo lineal, regresión simple 149
8.2 Métodos de ajuste de datos, métodos de mínimos cuadrados para ajustar una recta de regresión, propiedades de los estimadores 151
Análisis de varianza para regresión lineal 154
8.3 Intervalos de confianza 156
8.4 Pruebas de hipótesis 158
El análisis de varianza de la prueba *F* 159
Fuerza de la relación: el coeficiente de determinación 159
Interpretación de los resultados de una regresión significativa 160
8.5 Correlación 161
8.6 Predicción 164
8.7 Utilización de software de estadística para la solución de casos 168
Procedimientos de regresión lineal: *Microsoft Excel* 168
Procedimientos de regresión lineal: *MINITAB* 169
Ejercicios 172

9 REGRESIÓN MÚLTIPLE

9.1 Desarrollo del modelo de regresión múltiple 175
Significado de los coeficientes de regresión 176
9.2 El coeficiente de determinación múltiple, prueba *F* global 179
Análisis de varianza de la prueba *F* 180
9.3 Análisis de residuales 180
9.4 Inferencias respecto a los coeficientes de regresión de la población 181
Pruebas de hipótesis 181
Intervalos de confianza 182
9.5 El modelo de regresión cuadrática 182
9.6 Utilización de software de estadística para la solución de casos 185
Procedimientos de regresión múltiple: *Microsoft Excel* 185
Procedimientos de regresión múltiple: *MINITAB* 186
Ejercicios 188

10 SERIES DE TIEMPO

10.1 Componentes de la serie de tiempo 191
Variación horizontal 191
Variación de tendencia 193
Variación o componente estacional 194
Variaciones estacional y de tendencia 195

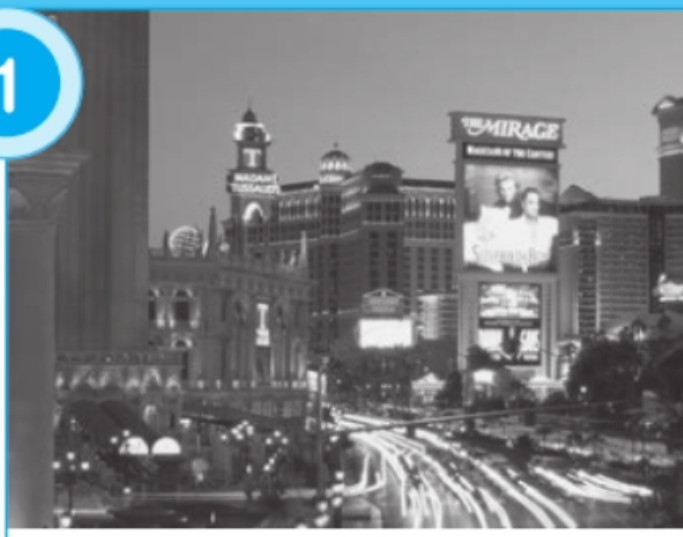
10.2 Análisis de tendencias 197
Regresión de tendencia lineal 197
Suavizamiento exponencial lineal de Holt 201
Regresión de tendencia no lineal 203
10.3 Variación estacional 205
Estacionalidad sin tendencia 205
Estacionalidad y tendencia 208
10.4 Patrón o componente cíclico 210
10.5 Variación irregular 210
Movimientos irregulares o al azar o ruido estadístico 211
10.6 Problemas que incluyen los cuatro componentes 211
10.7 Análisis de las series de tiempo en los pronósticos 215
Cálculo de los índices estacionales 216
Desestacionalización de una serie de tiempo 219
Uso de una serie de tiempo desestacionalizada para identificar tendencias 220
Ajustes estacionales 221
10.8 Utilización de software de estadística para la solución de casos 222
Elaboración de pronósticos con *MINITAB* 222
Ejercicios 225

11 ESTADÍSTICA NO PARAMÉTRICA
11.1 Introducción, pruebas no paramétricas 228
11.2 Prueba de los signos 228
Aproximación normal para la prueba del signo 230
11.3 Prueba de la suma de rangos 231
Aproximación normal para la prueba de suma de rango de Wilcoxon 235
11.4 Prueba de corridas con una muestra 237
Aproximación normal para la prueba de rango con signo de Wilcoxon 240
11.5 Correlación por rangos 241
11.6 Utilización de software de estadística para la solución de casos 245
Procedimientos no paramétricos: *MINITAB* 245
Ejercicios 248

APÉNDICE 251

1

Teoría del muestreo



© PictureNet/CORBIS

OBJETIVOS GENERALES

En este capítulo, empezamos a estudiar *muestras* y las *estadísticas* que las describen. Estas estadísticas muestrales se usan para hacer inferencias acerca de los correspondientes parámetros de población.

ÍNDICE DEL CAPÍTULO

- Muestras aleatorias
- Planes muestrales y diseños experimentales

NECESITO SABER...

Cuándo el tamaño de la muestra es lo bastante grande para usar el teorema del límite central
Cómo calcular probabilidades para la media muestral $\bar{x}$
Cómo calcular probabilidades para la proporción muestral $\hat{p}$

1.1 CONCEPTO DE MUESTRA

Una muestra es una pequeña porción de algo, representativa de un todo, que es usada para conocimiento público o para analizarla.

En estadística las muestras son tomas parciales de un conjunto total de cosas materiales o inmateriales que se quieren analizar, y que por su cantidad es imposible hacerlo sobre la cantidad total. Del análisis total de elementos tomados en cuenta se hacen inferencias sobre el universo. Por ejemplo, si se desea evaluar la calidad educativa de las instituciones de un país, se aplican exámenes en algunos institutos de distintas regiones aplicando técnicas de muestreo que consideren los centros representativos de distintas calidades de enseñanza.

1.2 CARACTERÍSTICAS DE LA MUESTRA

TIP MI CONSEJO

Parámetro ⇔ Población
Estadística ⇔ Muestra

En niveles anteriores, usted ha aprendido ya acerca de distribuciones de probabilidad; por ejemplo, las distribuciones binomial y normal. La forma de la distribución normal está determinada por su media μ y su desviación estándar σ, mientras que la forma de la distribución binomial está determinada por p. Estas medidas numéricas descriptivas, llamadas **parámetros**, son necesarias para calcular la probabilidad de observar resultados muestrales.

En situaciones prácticas, usted puede decidir qué *tipo* de distribución de probabilidad usar como modelo, pero los valores de los *parámetros* que especifican su *forma exacta* se desconocen. A continuación veamos dos ejemplos:

- Un entrevistador está seguro de que las respuestas a sus preguntas "de acuerdo/en desacuerdo" seguirán una distribución binomial, pero se desconoce p, la proporción de quienes están "de acuerdo" de la población.
- Un agrónomo cree que la producción por acre de una variedad de trigo está distribuida normalmente en forma aproximada, pero se desconocen la media μ y desviación estándar σ de la producción.

En estos casos, debemos apoyarnos en la *muestra* para saber de estos parámetros. La proporción de quienes están "de acuerdo" en la muestra del entrevistador da información acerca del valor real de p. La media y desviación estándar de la muestra del agrónomo aproximan los valores reales de μ y de σ. Pero, si se desea que la muestra dé *información confiable* acerca de la población, ¡la muestra debe ser seleccionada en cierta forma!

1.3 TIPOS DE MUESTREO

Muestreo aleatorio

La forma en que una muestra se selecciona recibe el nombre de **plan muestral** o **diseño experimental** y determina la cantidad de información de la muestra. Saber el plan muestral empleado en una situación particular permitirá medir la confiabilidad o bondad de la inferencia.

El **muestreo aleatorio simple** es un plan muestral de uso común en el que cada muestra de tamaño n tiene la misma probabilidad de ser seleccionada. Por ejemplo, supongamos que se desea seleccionar una muestra de tamaño $n = 2$ de una población que contiene $N = 4$ objetos. Si los cuatro objetos están identificados por los símbolos x_1, x_2, x_3 y x_4, hay seis pares distintos que podrían seleccionarse, como se muestra en la tabla 1.1. Si la muestra de $n = 2$ observaciones se selecciona de modo que cada una de estas seis muestras tenga la misma probabilidad de selección, dada por 1/6, entonces la muestra resultante se denomina **muestra aleatoria simple** o únicamente **muestra aleatoria**.

Definición Si una muestra de n elementos se selecciona de entre una población de N elementos, usando un plan muestral en el que cada una de las posibles muestras tiene la misma probabilidad de selección, entonces se dice que el muestreo es **aleatorio** y la muestra resultante es una **muestra aleatoria simple**.

TABLA 1.1 **Formas de seleccionar una muestra de tamaño 2 de 4 objetos**

Muestra	Observaciones en la muestra
1	x_1, x_2
2	x_1, x_3
3	x_1, x_4
4	x_2, x_3
5	x_2, x_4
6	x_3, x_4

Un muestreo aleatorio perfecto es difícil de obtener en la práctica. Si el tamaño N de la población es pequeño se podría escribir cada uno de los N números en una ficha, mezclar las fichas y seleccionar una muestra de n fichas. Los números que seleccione corresponden a las n mediciones que aparecen en la muestra. Como este método no siempre es práctico, un método más sencillo y confiable utiliza **números aleatorios**, es decir, dígitos generados de modo que los valores de 0 a 9 se presentan al azar y con igual frecuencia. Estos números pueden ser generados por computadora o incluso aparecer en una calculadora científica. De manera opcional, la tabla 10 del Apéndice es una tabla de números aleatorios que se pueden usar para seleccionar una *muestra aleatoria.*

EJEMPLO 1.1 Una base de datos de computadora en una firma de abogados contiene archivos para $N = 1\,000$ clientes. La firma desea seleccionar $n = 5$ archivos para revisión. Seleccione una muestra aleatoria simple de cinco archivos de esta base de datos.

Solución Primero debe marcar cada archivo con un número de 1 a 1 000. Quizá los archivos se guarden alfabéticamente y la computadora ya ha asignado un número a cada uno. A continuación genere una secuencia de 10 números aleatorios de tres dígitos. Si usa la tabla 10 del Apéndice, seleccione un punto inicial aleatorio y use una parte de la tabla similar a la de la tabla 1.2. El punto inicial aleatorio asegura que usted no use la misma secuencia una y otra vez.

Los primeros tres dígitos de la tabla 1.2 indican el número del primer archivo a revisarse. El número aleatorio 001 corresponde al archivo #1 y el último archivo, #1000, corresponde al número aleatorio 000. Usando la tabla 1.2, se seleccionarían los cinco archivos numerados 155, 450, 32, 882 y 350 para revisión. De manera opcional, podría elegir leer en sentido horizontal las líneas y seleccionar los archivos 155, 350, 989, 450 y 369 para revisión.

TABLA 1.2 **Parte de una tabla de números aleatorios**

15574	35026	98924
45045	36933	28630
03225	78812	50856
88292	26053	21121

La situación descrita en el ejemplo 1.1 se denomina **estudio observacional**, porque los datos ya existían antes que usted decidiera *observar* o describir sus características. La mayoría de los estudios muestrales, en los que la información se capta con un cuestionario, caen en esta categoría. Las bases de datos de computadora hacen posible asignar números de identificación a cada elemento aun cuando la población sea grande y seleccionar una muestra aleatoria simple. Tenga cuidado al efectuar un estudio muestral y esté atento a estos problemas que se presentan con frecuencia:

- **Sin respuesta:** Usted ha seleccionado su muestra aleatoria y enviado sus cuestionarios, pero sólo 50% de los entrevistados devolvió sus cuestionarios. ¿Las respuestas que usted recibió son representativas de toda la población o están **sesgadas** porque sólo quienes eran particularmente tendenciosos en el tema fueron seleccionados para responder?
- **Cobertura demasiado baja:** Usted ha seleccionado su muestra aleatoria usando registros telefónicos como una base de datos. ¿La base de datos que usó sistemáticamente excluye ciertos segmentos de la población, quizas aquellos que usan sólo celulares o que sus números son privados?

- **Sesgo verbal:** Su cuestionario puede tener preguntas que son demasiado complicadas o tienden a confundir al lector. Posiblemente las preguntas son sensibles por naturaleza, por ejemplo: "¿Alguna vez ha consumido usted drogas?" o "¿Alguna vez ha engañado en su declaración de impuestos?", y quienes responden no contestan con la verdad.

Se han diseñado métodos para resolver algunos de estos problemas, pero sólo si usted sabe que existen. Si su encuesta está *sesgada* por cualquiera de estos problemas, entonces sus conclusiones no serán muy confiables, ¡aunque haya seleccionado una muestra aleatoria!

Alguna investigación realizada comprende la **experimentación**, en la que una condición experimental o *tratamiento* se impone en las *unidades experimentales*. Seleccionar una muestra aleatoria simple es más difícil en esta situación.

EJEMPLO 1.2

Una química investigadora somete a prueba un nuevo método para medir la cantidad de titanio (Ti) en muestras de mineral. Selecciona 10 muestras de mineral del mismo peso para su experimento. Cinco de las muestras se medirán con un método estándar y las otras cinco con el nuevo método. Use números aleatorios para asignar las 10 muestras de mineral a los grupos nuevo y estándar. ¿Estos datos representan una muestra aleatoria simple de entre la población?

Solución Hay realmente dos poblaciones en este experimento. Están formadas por mediciones de titanio, usando ya sea el método nuevo o el estándar, para *todas las posibles* muestras de mineral de este peso. Estas poblaciones no existen en realidad; son **poblaciones hipotéticas**, imaginadas en la mente de la investigadora. Entonces, es imposible seleccionar una muestra aleatoria simple usando los métodos del ejemplo 1.1. En cambio, la investigadora selecciona lo que cree que son 10 muestras *representativas* de mineral y espera que éstas se *comportarán como si* se hubieran seleccionado al azar de las dos poblaciones.

La investigadora puede, no obstante, seleccionar al azar las cinco muestras a medir con cada método. Numere las muestras del 1 al 10. Las cinco muestras seleccionadas para el nuevo método pueden corresponder a cinco números aleatorios de un dígito. Use esa secuencia de dígitos aleatorios generados en una calculadora científica:

948247817184610

Como no se puede seleccionar dos veces la misma muestra de mineral hay que saltarse cualquier dígito que ya haya sido elegido. Las muestras 9, 4, 8, 2 y 7 se medirán usando el nuevo método. Las otras muestras, es decir 1, 3, 5, 6 y 10, se medirán con el método estándar.

Muestreo estratégico

Además del *muestreo aleatorio simple*, hay otros planes muestrales con carácter aleatorio y, por tanto, dan una base probabilística para hacer inferencias. Tres de esos planes están basados en *muestreo estratificado, conglomerado* y *sistemático.*

Cuando la población está formada por dos o más subpoblaciones, llamadas **estratos**, un plan muestral que asegura que cada subpoblación está representada en la muestra se denomina **muestra aleatoria estratificada**.

Definición Un **muestreo aleatorio estratificado** comprende seleccionar una muestra aleatoria simple de cada uno de un número dado de subpoblaciones o **estratos**.

Las opiniones de la ciudadanía acerca de la construcción de un centro de artes interpretativas podrían ser recolectadas usando una muestra aleatoria estratificada con distritos de votación como estratos. Los sondeos nacionales comprenden alguna forma muestral aleatoria estratificada con los estados como estratos.

Muestreo sistemático

En ocasiones la población que será muestreada está ordenada, por ejemplo, una lista alfabetizada de personas con licencias de manejo, una lista de usuarios de la compañía de luz ordenada por direcciones de servicio o una lista de clientes por números de cuenta. En estas y otras situaciones, se elige al azar un elemento de los primeros k elementos y, a continuación, cada k-ésimo elemento de ahí en adelante se incluye en la muestra.

Definición Una **muestra aleatoria sistemática 1 en *k*** involucra la selección aleatoria de uno de los primeros k elementos de una población ordenada y luego la selección sistemática de cada k-ésimo elemento de ahí en adelante.

TIP MI CONSEJO

Todos los planes muestrales empleados para hacer inferencias comprenden el *proceso aleatorio*

Sin embargo, no todos los planes muestrales comprenden una selección aleatoria. Es probable que usted haya oído de las encuestas telefónicas no aleatorias, en las que las personas que desean expresar apoyo a una pregunta llaman a un "número 900" y los que se oponen llaman a un segundo "número 900". Cada persona debe pagar por su llamada. Es obvio que quienes llaman no representan la población en general. Este tipo muestral es una forma de una **muestra de conveniencia**, es decir, una muestra que se puede obtener de manera fácil y sencilla sin selección aleatoria. Hacer publicidad a personas a quienes se les pagará una cuota por participar en un experimento produce una muestra de conveniencia. El **muestreo de juicio** permite que la persona que haga el muestreo decida quién estará o no incluido en la muestra. El **muestreo de cuota**, en el que la composición de la muestra debe reflejar la composición de la población en alguna característica preseleccionada, con frecuencia tiene un componente no aleatorio en el proceso de selección. **Recuerde que las muestras no aleatorias se pueden describir pero no deben usarse para hacer inferencias**.

Muestreo de conglomerados

Se utiliza otra forma muestral aleatoria cuando las unidades muestrales disponibles son grupos de elementos, llamados **conglomerados**. Por ejemplo, una familia es un *conglomerado* de personas que viven juntas. Una manzana o vecindario de una ciudad podría ser una unidad muestral conveniente y ser considerada un *conglomerado* para un plan muestral determinado.

Definición Una **muestra de conglomerados** es una muestra aleatoria simple tomada de los conglomerados disponibles en la población.

Cuando un conglomerado particular se incluye en la muestra, se toma un censo de cada uno de los elementos del conglomerado.

EJERCICIOS

TÉCNICAS BÁSICAS

1.1 Una población está formada por $N = 500$ unidades experimentales. Use una tabla numérica aleatoria para seleccionar una muestra aleatoria de $n = 20$ unidades experimentales. (AYUDA: Como es necesario usar números de tres dígitos, asigne dos números de tres dígitos a cada una de las unidades muestrales en la forma ilustrada en la tabla.) ¿Cuál es la probabilidad de que cada unidad experimental sea seleccionada para su inclusión en la muestra?

Unidades experimentales	Números aleatorios
1	001, 501
2	002, 502
3	003, 503
4	004, 504
.	.
.	.
.	.
499	499, 999
500	500, 000

1.2 Un analista político desea seleccionar una muestra de $n = 20$ personas de una población de 2 000. Use la tabla numérica aleatoria para identificar a las personas que se incluirán en la muestra.

1.3 Una población contiene 50 000 votantes. Use la tabla numérica aleatoria para identificar a los votantes que se incluirán en una muestra aleatoria de $n = 15$.

1.4 Una pequeña ciudad contiene 20 000 votantes. Use la tabla numérica aleatoria para identificar a los votantes que se incluirán en una muestra aleatoria de $n = 15$.

APLICACIONES

1.5 Cada décima persona Se obtuvo una muestra aleatoria de opinión pública, en una pequeña ciudad, al seleccionar cada décima persona que pasara por la esquina de mayor movimiento en el centro de la ciudad. ¿Esta muestra tendrá las características de una muestra aleatoria seleccionada de los ciudadanos? Explique.

1.6 Parques y recreación Se envió por correo un cuestionario a 1 000 votantes municipales registrados seleccionados al azar. Sólo 500 cuestionarios fueron devueltos y, de éstos, 360 contestaron que se oponían rotundamente a un recargo propuesto para sostener al Departamento de Parques y Recreación de la ciudad. ¿Está usted dispuesto a aceptar la cifra de 72% como estimación válida del porcentaje en la ciudad que se oponen al recargo? ¿Por qué?

1.7 Listas del DMV En muchos estados, unas listas de posibles miembros de jurados se conforman a partir de listas de registro de votantes, así como de registros del Departamento de Vehículos a Motor de conductores con licencia de manejo y propietarios de autos. ¿En qué formas podría esta lista no cubrir adecuadamente ciertos sectores de la población?

1.8 Sexo y violencia Una pregunta en un cuestionario de cierta encuesta dice: "¿No está usted de acuerdo en que hay demasiado sexo y violencia durante el horario estelar en televisión?" Comente sobre posibles problemas con las respuestas a esta pregunta. Sugiera una mejor forma de plantear la pregunta.

1.9 Ácidos grasos omega 3 Contrario a la opinión actual sobre los ácidos grasos omega 3, una nueva investigación muestra que las grasas benéficas quizá no ayuden a reducir segundos infartos en los sobrevivientes que han sufrido uno. El estudio incluyó 4 837 hombres y mujeres que recibían tratamiento por enfermedad cardiaca. El grupo experimental recibió 400 mg adicionales diarios de las grasas.[1] Suponga que este experimento se repitió con 50 individuos en el grupo de control y 50 individuos en el grupo experimental. Determine un esquema de aleatorización para asignar los 100 individuos a los dos grupos.

1.10 ¿Discriminación racial? ¿Es importante la raza de un entrevistador? Esta pregunta fue investigada por Chris Gilberg y colegas, y publicada en una edición de la revista *Chance*.[2] El entrevistador preguntaba: "¿Piensa usted que la acción afirmativa debe usarse como criterio de selección de ocupación?" con posibles respuestas de sí o no.

a. ¿Qué problemas esperaría usted con respuestas a esta pregunta cuando las hacen entrevistadores de diferentes orígenes étnicos?

b. Cuando las personas eran entrevistadas por un afroestadounidense, la respuesta era alrededor de 70% a favor de una acción afirmativa, aproximadamente 35% cuando eran entrevistadas por un asiático y alrededor de 25% cuando eran entrevistadas por un caucásico. ¿Estos resultados apoyan su respuesta de la parte **a**?

1.11 Juventud de estadounidenses nativos *American Journal of Human Biology* publicó un estudio sobre una herramienta de evaluación de la dieta para usarla en la población de jóvenes nativos estadounidenses urbanos.[3] Los sujetos eran jóvenes estadounidenses nativos que estudiaban en un programa después de clase en Minneapolis, MN. Los 61 niños entre 9 y 13 años de edad que cumplieron con los requisitos de los objetivos de estudio fueron incluidos en el experimento.

a. Describa el plan muestral empleado para seleccionar participantes para el estudio.

b. ¿Qué mecanismo de probabilidad se utilizó para seleccionar esta muestra de 61 nativos estadounidenses de 9 a 13 años de edad?

c. ¿Se pueden hacer inferencias válidas usando los resultados de este estudio? ¿Por qué?

d. Si usted tuviera que idear un plan muestral alternativo, ¿qué cambiaría?

1.12 Tai chi y fibromialgia Un nuevo estudio pequeño muestra que el tai chi, una práctica china antigua de ejercicio y meditación, puede aliviar síntomas de fibromialgia crónica dolorosa. El estudio asignó 66 pacientes con fibromialgia para que tomaran una clase de tai chi durante 12 semanas, o asistieran a una clase de educación para la salud.[4]

a. Proporcione un esquema de aleatorización para asignar los 66 sujetos a los dos grupos.

b. ¿Su esquema de aleatorización producirá grupos de tamaño igual? Explique.

1.13 Ir a la Luna Se llevaron a cabo dos encuestas Gallup diferentes para *CNN/USA Today*, donde aparecía lo que pensaba la sociedad acerca del programa espacial de Estados Unidos.[5] Veamos a continuación una pregunta de cada encuesta, junto con las respuestas de los estadounidenses encuestados:

Exploración espacial

Encuesta Gallup/CNN/USA Today: Nacional:

"¿Estaría usted a favor o en contra de un nuevo programa espacial de Estados Unidos para enviar astronautas a la Luna?", Forma A (N = 510, MoE ± 5)

	A favor	En contra	No tiene opinión
	%	%	%
12/03	53	45	2

"¿Estaría usted a favor o en contra de que el gobierno gaste miles de millones de dólares para enviar astronautas a la Luna?", Forma B (N = 494, MoE ± 5)

	A favor	En contra	No tiene opinión
	%	%	%
12/03	31	67	2

a. Lea las dos preguntas de la encuesta. ¿Cuál de las dos redacciones es más imparcial? Explique.

b. Vea las respuestas para las dos encuestas. ¿Cómo explicaría usted las grandes diferencias en los porcentajes, ya sea a favor o en contra del nuevo programa?

1.14 Estados Unidos, ¡pregunta! Una encuesta de 2003 de la política nacional, titulada "Estados Unidos, ¡pregunta!" fue enviada por la Comisión Nacional Republicana del Congreso a votantes del 44 Distrito del Congreso, pidiendo opiniones en varios asuntos políticos.[6] A continuación veamos algunas preguntas de la encuesta:

- En años recientes, ¿el gobierno federal ha intervenido más, o menos, en sus asuntos personales o profesionales?
- ¿Tiene razón el presidente Bush al tratar de refrenar el tamaño y alcance del gobierno federal, contra los deseos de los grandes demócratas en el gobierno?
- ¿Piensa que la pena de muerte es un factor disuasivo para la delincuencia?
- ¿Está usted de acuerdo en que a los demócratas obstruccionistas no se les permita obtener el control del Congreso de Estados Unidos en las próximas elecciones?

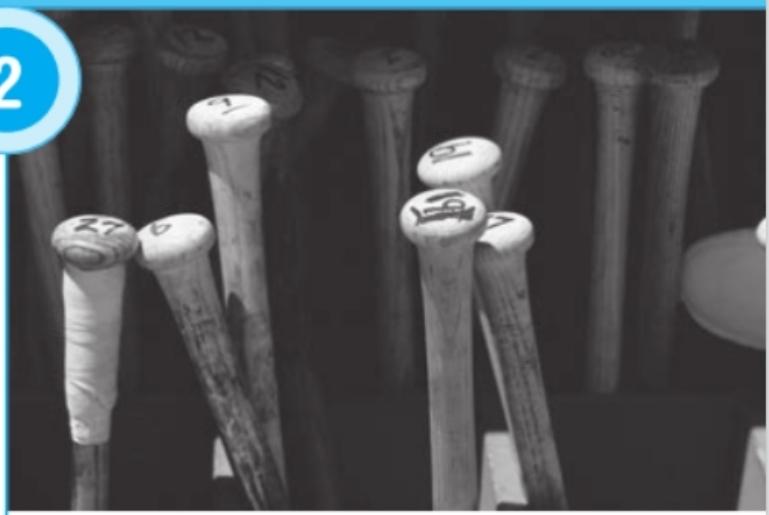
Comente sobre el efecto del sesgo en la redacción sobre las respuestas reunidas usando esta encuesta.

Fuentes

1. Alice Park, "Omega-3 May Reduce Heart Risks Less Than Thought." http://www.time.com/time/health/article/0,8599,2014603,00.html, 30 de agosto de 2010.
2. Chris Gilberg, J.L. Cos, H. Kashima y K. Eberle, "Survey Biases: When Does the Interviewer's Race Matter?" *Chance*, otoño de 1996, p. 23.
3. Chery Smith y Stefanie Fila, "Comparison of the Kid's Block Food Frequency Questionnaire to the 24-Hour Recall in Urban Native American Youth," *American Journal of Human Biology*, 18:706–709 (2006).
4. Liz Szabo,"Study: Tai Chi Could Ease Fibromyalgia Pain," *The Press-Enterprise* (Riverside, CA), 19 de agosto de 2010, p. 4D.
5. "Space Exploration," CNN/*USA Today*/Gallup Poll, http://www.pollingreport.com/science.htm#Space, 5 de abril de 2004.
6. "ASK AMERICA: 2003 Nationwide Policy Survey," Congressional District #44, 23 de junio de 2003.

2

Distribuciones muestrales



© Joe Sohm-VisionsofAmerica/Photodisc/Getty

OBJETIVOS GENERALES

En el capítulo anterior estudiamos *poblaciones* y los *parámetros* que las describen. Estas poblaciones eran discretas o continuas y empleamos la *probabilidad* como herramienta para determinar qué tan probables podrían ser ciertos resultados muestrales. Este capítulo comprende muestreo y distribuciones muestrales, que describen el comportamiento de estadísticas muestrales en muestreo repetido.

ÍNDICE DEL CAPÍTULO

- Control de un proceso estadístico: gráficas $\bar{x}$
- El teorema del límite central
- Estadísticas y distribuciones muestrales
- La distribución muestral de la media muestral, $\bar{x}$
- La distribución muestral de la proporción muestral, $\hat{p}$

NECESITO SABER...

Cuándo el tamaño de la muestra es lo bastante grande para usar el teorema del límite central
Cómo calcular probabilidades para la media muestral $\bar{x}$
Cómo calcular probabilidades para la proporción muestral $\hat{p}$

2.1 CONCEPTOS FUNDAMENTALES

Cuando se selecciona una muestra aleatoria de una población las medidas numéricas descriptivas que se calculen de la muestra se denominan **estadísticas**. Las estadísticas varían o cambian para cada muestra aleatoria diferente que se seleccione; esto es, son *variables aleatorias*. Las distribuciones de probabilidad para estadísticas se llaman **distribuciones muestrales** porque, en muestreos repetidos, dan esta información:

- Qué valores de la estadística pueden presentarse.
- Con qué frecuencia se presenta cada valor.

Definición La **distribución muestral de una estadística** es la distribución de probabilidad para los posibles valores de la estadística, que resulta cuando muestras aleatorias de tamaño n se sacan repetidamente de la población.

Hay tres formas de hallar la distribución muestral de una estadística:

1. Deducir la distribución *matemáticamente* usando las leyes de probabilidad.
2. Usar una *simulación* para aproximar la distribución. Esto es, sacar un gran número de muestras de tamaño n, calculando el valor de la estadística para cada muestra y tabular los resultados en un histograma de frecuencia relativa. Cuando el número de muestras es grande, el histograma será muy cercano a la distribución muestral teórica.
3. Usar *teoremas estadísticos* para obtener distribuciones muestrales exactas o aproximadas.

El siguiente ejemplo demuestra cómo deducir las distribuciones muestrales de dos estadísticas para una población muy pequeña.

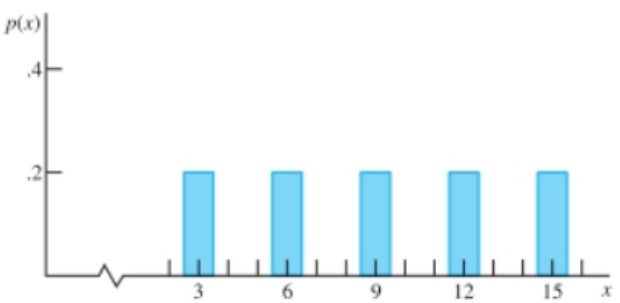
EJEMPLO 2.1

Una población está formada por $N = 5$ números: 3, 6, 9, 12, 15. Si una muestra aleatoria de tamaño $n = 3$ se selecciona sin reemplazo, encuentre las distribuciones muestrales para la media muestral $\bar{x}$ y la mediana m.

Solución Usted está muestreando de la población que aparece en la figura 2.1. Contiene cinco números distintos y cada uno es igualmente probable, con probabilidad $p(x) = 1/5$. Con facilidad puede hallar la media poblacional y mediana como

$$\mu = \frac{3 + 6 + 9 + 12 + 15}{5} = 9 \quad \text{y} \quad M = 9$$

FIGURA 2.1
Histograma de probabilidad para los $N = 5$ valores poblacionales del ejemplo 1.3



TIP MI CONSEJO
Las distribuciones muestrales pueden ser discretas o continuas

Hay 10 posibles muestras aleatorias de tamaño $n = 3$ y cada una es igualmente probable, con probabilidad 1/10. Estas muestras, junto con los valores calculados de $\bar{x}$ y m para cada una, se muestran en la tabla 2.1. Usted observará que algunos valores de $\bar{x}$ son más probables que otros porque se presentan en más de una muestra. Por ejemplo,

$$P(\bar{x} = 8) = \frac{2}{10} = .2 \quad \text{y} \quad P(m = 6) = \frac{3}{10} = .3$$

Los valores de la tabla 2.1 están tabulados, y las distribuciones muestrales para $\bar{x}$ y m se muestran en la tabla 2.2 y la figura 2.2.

Como la población de $N = 5$ valores es simétrica cerca del valor de $x = 9$, la *media poblacional* y la *mediana* son iguales a 9. Parecería razonable, por tanto, considerar usar ya sea $\bar{x}$ o m como estimador posible de $M = \mu = 9$. ¿Cuál estimador elegiría usted? De la tabla 2.2 se observa que, al usar m como estimador, estaría en un error de $9 - 6 = 3$ con probabilidad .3 o de $9 - 12 = -3$ con probabilidad .3. Esto es, el error en estimación usando m sería 3 con probabilidad .6. Al usar $\bar{x}$, no obstante, un error de 3 ocurriría con probabilidad de sólo .2. Únicamente en estos casos, podría usarse $\bar{x}$ como estimador en preferencia sobre m.

TABLA 2.1 **Valores de $\bar{x}$ y m para muestreo aleatorio simple cuando $n = 3$ y $N = 5$**

Muestra	Valores muestrales	$\bar{x}$	m
1	3, 6, 9	6	6
2	3, 6, 12	7	6
3	3, 6, 15	8	6
4	3, 9, 12	8	9
5	3, 9, 15	9	9
6	3, 12, 15	10	12
7	6, 9, 12	9	9
8	6, 9, 15	10	9
9	6, 12, 15	11	12
10	9, 12, 15	12	12

TABLA 2.2 **Distribuciones muestrales para a) la media muestral y b) la mediana muestral**

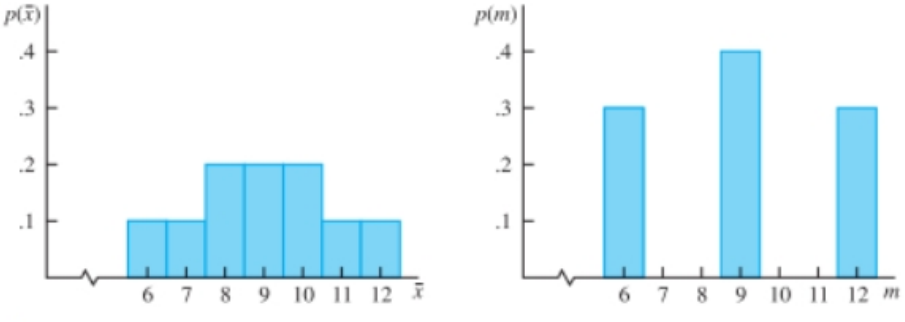
(a)

$\bar{x}$	$p(\bar{x})$
6	.1
7	.1
8	.2
9	.2
10	.2
11	.1
12	.1

(b)

m	$p(m)$
6	.3
9	.4
12	.3

FIGURA 2.2
Histogramas de probabilidad para las distribuciones muestrales de la media muestral, $\bar{x}$, y la mediana muestral, m, del ejemplo 1.3



TIP **MI CONSEJO**

Casi toda estadística tiene una media y una desviación estándar (o *error estándar*) que describen su centro y dispersión

No fue difícil deducir estas distribuciones muestrales en el ejemplo 2.1 porque el número de elementos en la población era muy pequeño. Cuando éste no es el caso, es necesario usar uno de estos métodos:

- Use una simulación para aproximar la distribución muestral de manera empírica.
- Apóyese en teoremas estadísticos y resultados teóricos.

Un teorema estadístico importante, que describe la distribución muestral de estadísticas que son sumas o promedios, se presenta en la siguiente sección.

2.2 DISTRIBUCIÓN MUESTRAL DE LA MEDIA, PROPIEDADES, ERROR ESTÁNDAR DE LA MEDIA

Si la media poblacional μ es desconocida se pueden seleccionar varias *estadísticas* como estimador; la media muestral $\bar{x}$ y la mediana muestral m son dos que con facilidad llegan a nuestra mente. ¿Cuál debe usarse? Considere estos criterios al seleccionar el estimador para μ:

- ¿Es fácil o difícil de calcular?
- ¿Produce estimaciones que sean demasiado altas o demasiado bajas de manera consistente?
- ¿Es más o menos variable que otros estimadores posibles?

Las distribuciones muestrales para $\bar{x}$ y m con $n = 3$ para la pequeña población del ejemplo 2.1 mostraron que, en términos de estos criterios, la media muestral funcionó mejor que la mediana muestral como estimador de μ. En muchas situaciones, la media muestral $\bar{x}$ tiene propiedades deseables como un estimador que no son compartidas por otros estimadores competidores; por tanto, se emplea en forma más amplia.

LA DISTRIBUCIÓN MUESTRAL DE LA MEDIA MUESTRAL, $\bar{x}$

- Si una muestra aleatoria de n mediciones se selecciona de una población con media μ y desviación estándar μ, la distribución muestral de la media muestral $\bar{x}$ tendrá media μ y desviación estándar*

$$\frac{\sigma}{\sqrt{n}}$$

- Si la población tiene una distribución normal, la distribución muestral de $\bar{x}$ estará exactamente *distribuida* en forma normal, *cualquiera que sea el tamaño muestral n*.
- Si la distribución poblacional *no es normal*, la distribución muestral de $\bar{x}$ estará distribuida normalmente en *forma aproximada* para muestras grandes (por el teorema del límite central). De manera conservadora, se requiere $n \geq 30$.

* Cuando muestras repetidas de tamaño n se seleccionan al azar de una población *finita* con N elementos cuya media sea μ y varianza σ^2, la desviación estándar de $\bar{x}$ es

$$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

donde σ^2 es la varianza poblacional. Cuando N es grande con respecto al tamaño muestral n, $\sqrt{(N-n)(N-1)}$ es aproximadamente igual a 1, y la desviación estándar de $\bar{x}$ es

$$\frac{\sigma}{\sqrt{n}}$$

Error estándar

Definición La desviación estándar de una estadística empleada como estimador de un parámetro poblacional también se denomina **error estándar del estimador** (abreviado **SE**) porque se refiere a la precisión del estimador. Por tanto, la desviación estándar de $\bar{x}$, dada por $\sigma/\sqrt{n}$, se conoce como **error estándar de la media** (abreviada SE($\bar{x}$) o sólo SE).

NECESITO SABER...

Cómo calcular probabilidades para la media muestral $\bar{x}$

Si se sabe que la distribución muestral de $\bar{x}$ es *normal o aproximadamente normal*, se puede describir el comportamiento de la media muestral $\bar{x}$ al calcular la probabilidad de observar ciertos valores de $\bar{x}$ en muestreo repetido.

1. Encuentre μ y calcule SE $(\bar{x}) = \sigma/\sqrt{n}$.
2. Escriba el evento de interés en términos de $\bar{x}$ y localice el área apropiada en la curva normal.
3. Convierta los valores necesarios de $\bar{x}$ en valores z usando
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$
4. Use la tabla 3 del Apéndice para calcular la probabilidad.

EJEMPLO 2.2 La duración de la enfermedad de Alzheimer desde el principio de los síntomas hasta el fallecimiento varía de 3 a 20 años; el promedio es 8 años con una desviación estándar de 4 años. El administrador de un gran centro médico selecciona al azar los registros médicos de 30 pacientes de Alzheimer ya fallecidos, de la base de datos del centro médico y anota la duración promedio. Encuentre las probabilidades aproximadas para estos eventos:

1. La duración promedio es menor a 7 años.
2. La duración promedio excede de 7 años.
3. La duración promedio está a no más de 1 año de la media poblacional $\mu = 8$.

Solución Como el administrador ha seleccionado una muestra aleatoria de la base de datos en este centro médico, puede sacar conclusiones acerca de pacientes pasados, presentes o futuros con enfermedad de Alzheimer en este centro médico. Si, por el contrario, este centro médico puede ser considerado como representativo de otros centros médicos del país, es posible sacar conclusiones de más largo alcance.

TIP MI CONSEJO

Si x es normal, $\bar{x}$ es normal para cualquier n

Si x no es normal, $\bar{x}$ es aproximadamente normal para n grande

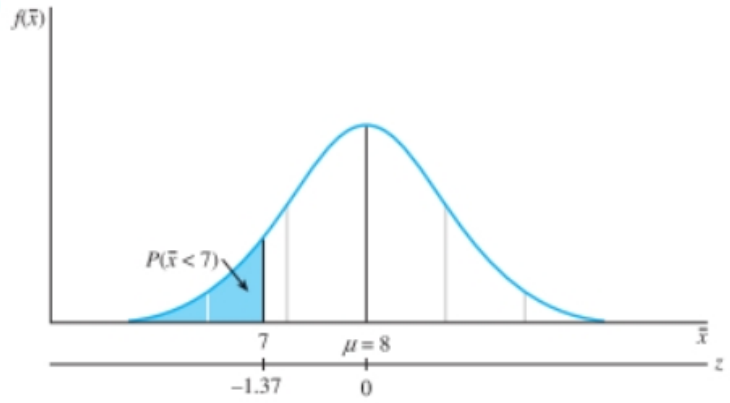
¿Qué se puede decir acerca de la forma de la población muestreada? No es simétrica, porque la media $\mu = 8$ no está a la mitad entre los valores máximo y mínimo. Como la media está más cerca al valor mínimo, la distribución está sesgada a la derecha, con unos pocos pacientes viviendo largo tiempo después que inicia la enfermedad. Cualquiera que sea la forma de la distribución poblacional, no obstante, la distribución muestral de $\bar{x}$ tiene una media de $\mu = 8$ y desviación estándar $\sigma/\sqrt{n} = 4/\sqrt{30} = .73$. Además, como el tamaño muestral es $n = 30$, el teorema del límite central asegura la *normalidad aproximada* de la distribución muestral de $\bar{x}$.

MI APPLET EN LÍNEA

Probabilidades normales para medias

1. La probabilidad de que $\bar{x}$ sea menor que 7 está dada por el área sombreada de la figura 2.3. Para hallar esta área, es necesario calcular el valor de z correspondiente a $\bar{x} = 7$:
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{7 - 8}{.73} = -1.37$$
De la tabla 3 del Apéndice se puede hallar el área acumulativa correspondiente a $z = -1.37$ y
$$P(\bar{x} < 7) = P(z < -1.37) = .0853$$

FIGURA 2.3
Probabilidad de que $\bar{x}$ sea menor que 7 para el ejemplo 1.4



(NOTA: Se debe usar $\sigma/\sqrt{n}$ (no σ) en la fórmula para z porque estamos buscando un área bajo la distribución muestral para $\bar{x}$, no bajo la distribución de probabilidad para x.)

TIP MI CONSEJO

Recuerde que para variables aleatorias continuas, no hay probabilidad asignada a un punto solo. Por tanto, $P(\bar{x} \leq 7) = P(\bar{x} < 7)$

2. El evento de que $\bar{x}$ exceda de 7 es el complemento del evento que $\bar{x}$ sea menor que 7. Entonces, la probabilidad de que $\bar{x}$ exceda de 7 es

$$P(\bar{x} > 7) = 1 - P(\bar{x} \leq 7)$$
$$= 1 - .0853 = .9147$$

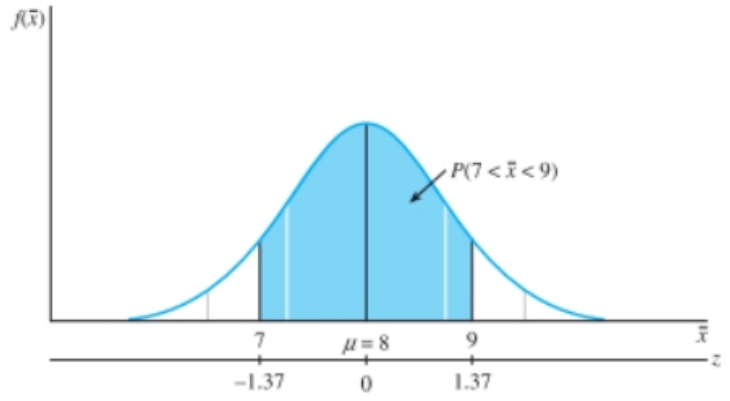
3. La probabilidad de que $\bar{x}$ se encuentre a no más de 1 año de $\mu = 8$ es el área sombreada en la figura 2.4. El valor z correspondiente a $\bar{x} = 7$ es $z = -1.37$, de la parte 1 y el valor z para $\bar{x} = 9$ es

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{9 - 8}{.73} = 1.37$$

La probabilidad de interés es

$$P(7 < \bar{x} < 9) = P(-1.37 < z < 1.37)$$
$$= .9147 - .0853 = .8294$$

FIGURA 2.4
La probabilidad de que $\bar{x}$ se encuentre a no más de 1 año de $\mu = 8$ para el ejemplo 2.2



EJEMPLO 2.3

Para evitar dificultades con la Comisión Federal de Comercio de Estados Unidos o con oficinas locales y estatales de protección al consumidor, un embotellador de refrescos debe estar razonablemente seguro de que sus botellas de 12 onzas en realidad contengan 12 onzas de líquido. Para determinar si una máquina embotelladora está funcionando de manera satisfac-

toria, un embotellador muestrea al azar 10 botellas por hora y mide la cantidad de líquido de cada botella. La media $\bar{x}$ de las 10 medidas de llenado se usa para determinar si se reajusta la cantidad de líquido introducido en la botella por la máquina llenadora.

Si los registros muestran que la cantidad de líquido por botella está normalmente distribuida, con una desviación estándar de .2 onzas y si la máquina embotelladora está ajustada para producir un llenado medio por botella de 12.1 onzas, ¿cuál es la probabilidad aproximada de que la media muestral $\bar{x}$ de las 10 botellas sea menor que 12 onzas?

Solución La media de la distribución muestral de la media muestral $\bar{x}$ es idéntica a la media de la población de llenados de botella, es decir, $\mu = 12.1$ onzas y el error estándar (SE) de $\bar{x}$ es

$$\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{.2}{\sqrt{10}} = .063$$

(NOTA: σ es la desviación estándar de la población de llenados de botella y n es el número de botellas de la muestra.) Como la cantidad de llenado está distribuida normalmente, $\bar{x}$ también está distribuida normalmente, como se muestra en la figura 2.5.

Para hallar la probabilidad de que $\bar{x}$ sea menor que 12 onzas, exprese el valor $\bar{x} = 12$ en unidades de desviación estándar:

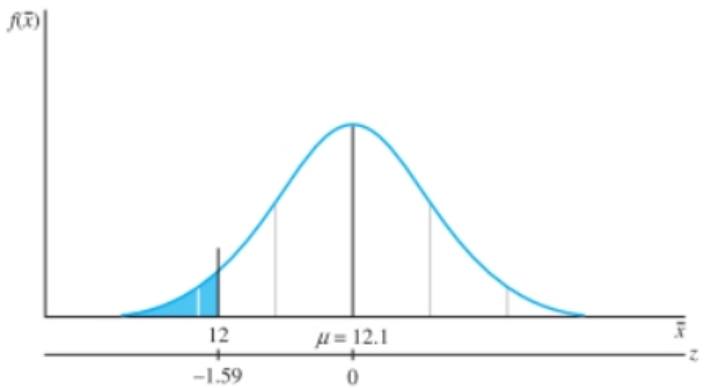
$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{12 - 12.1}{.063} = -1.59$$

Entonces

$$P(\bar{x} < 12) = P(z < -1.59) = .0559 \approx .056$$

Por tanto, si la máquina está ajustada para producir un llenado promedio de 12.1 onzas, el llenado medio $\bar{x}$ de una muestra de 10 botellas será menor a 12 onzas con una probabilidad igual a .056. Cuando se presenta esta señal de riesgo ($\bar{x}$ es menor que 12), el embotellador toma una muestra más grande para volver a verificar el ajuste de la máquina llenadora.

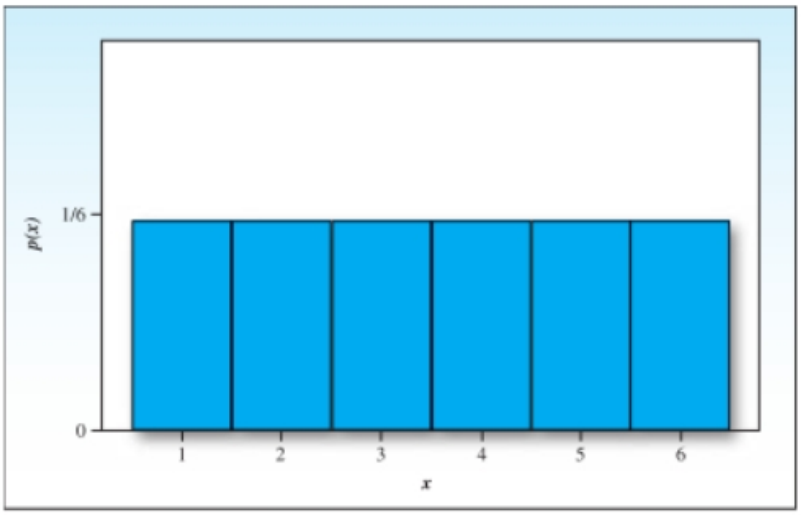
FIGURA 2.5
Distribución de probabilidad de $\bar{x}$, la media de los $n = 10$ llenados de botellas, para el ejemplo 2.3



2.3 TEOREMA DEL LÍMITE CENTRAL

El **teorema del límite central** establece que, en condiciones más bien generales, las sumas y medias de muestras aleatorias de mediciones tomadas de una población tienden a tener una distribución aproximadamente normal. Suponga que se tira un dado balanceado $n = 1$ vez. La variable aleatoria x es el número observado en la cara superior. Esta variable aleatoria conocida puede tomar seis valores, cada uno con probabilidad 1/6 y su distribución de probabilidad se muestra en la figura 2.6. La forma de la distribución es *plana* o *uniforme* y simétrica alrededor de la media $\mu = 3.5$, con una desviación estándar $\sigma = 1.71$.

FIGURA 2.6
Distribución de probabilidad para x, el número que aparece en un solo tiro de un dado

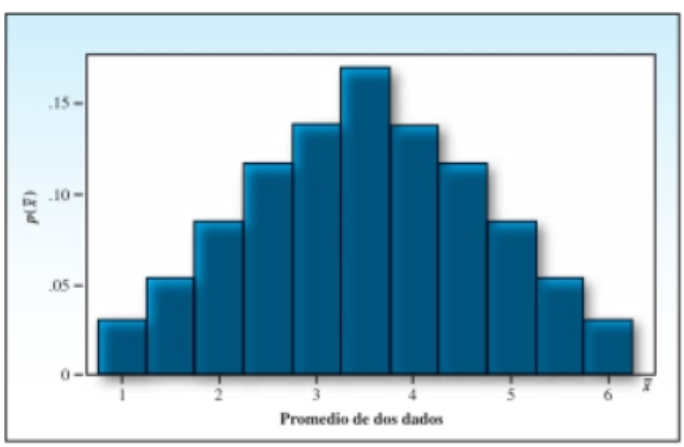


Ahora, tome un tamaño muestral de $n = 2$ de esta población; esto es, tire dos veces el dado y registre la suma de los números de las dos caras superiores, $\Sigma x_i = x_1 + x_2$. La tabla 2.3 muestra los 36 posibles resultados, cada uno con probabilidad 1/36. Las sumas están tabuladas y cada una de las posibles sumas está dividida entre $n = 2$ para obtener un promedio. El resultado es la **distribución muestral** de $\bar{x} = \Sigma x_i/n$, que se ilustra en la figura 2.7. Usted debe observar la considerable diferencia en la forma de la distribución muestral. Ahora tiene casi forma de montículo pero todavía es simétrica alrededor de la media $\mu = 3.5$.

TABLA 2.3 **Sumas de las caras superiores de dos dados**

	Primer dado					
Segundo dado	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

FIGURA 2.7
Distribución muestral de $\bar{x}$ para $n = 2$ dados



APPLET EN LÍNEA
Teorema del límite central

Usando un procedimiento similar generamos las distribuciones muestrales de $\bar{x}$ cuando $n = 3$ y $n = 4$. Para $n = 3$, la distribución muestral de la figura 2.8 con toda claridad muestra la forma de montículo de la distribución normal de probabilidad, todavía centrada en $\mu = 3.5$.

Observe también que la dispersión de la distribución es lentamente *decreciente* cuando *aumenta* el tamaño muestral n. La figura 2.9 muestra de forma espectacular que la distribución de $\bar{x}$ está distribuida normalmente en forma aproximada con base en una muestra de sólo $n = 4$. Este fenómeno es el resultado de un importante teorema estadístico llamado **teorema del límite central (TLC)**.

FIGURA 2.8
Distribución muestral de $\bar{x}$ para $n = 3$ dados

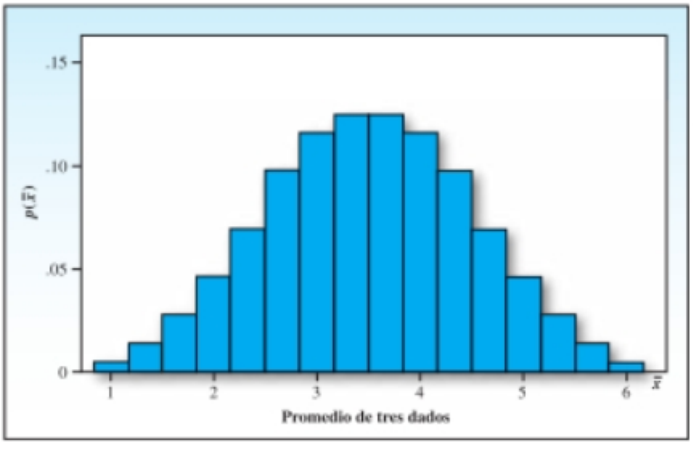
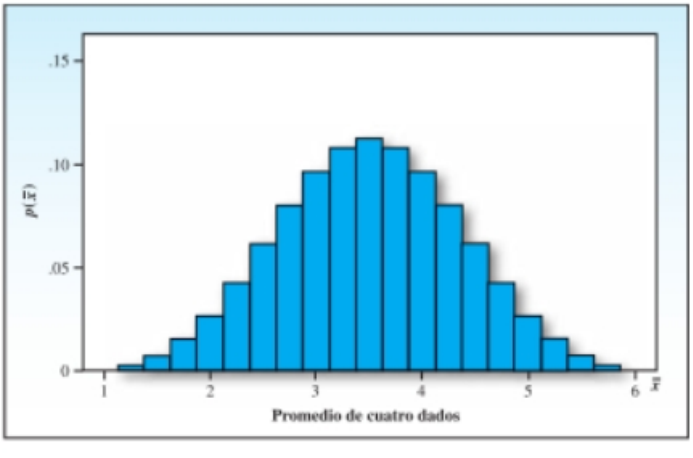


FIGURA 2.9
Distribución muestral de $\bar{x}$ para $n = 4$ dados



Teorema del límite central

Si muestras aleatorias de n observaciones se sacan de una población no normal con media finita μ y desviación estándar σ, entonces, cuando n es grande, la distribución de muestreo de la media muestral $\bar{x}$ está distribuida normalmente en forma aproximada, con media μ y desviación estándar

$$\frac{\sigma}{\sqrt{n}}$$

La aproximación se hace más precisa cuando n se hace grande.

TIP MI CONSEJO

La distribución muestral de $\bar{x}$ siempre tiene una media μ y desviación estándar $\sigma/\sqrt{n}$. El teorema del límite central ayuda a describir su forma

Cualquiera que sea su forma, la distribución muestral de $\bar{x}$ siempre tiene una media idéntica a la media de la población muestreada y una desviación estándar igual a la desviación poblacional estándar σ dividida entre $\sqrt{n}$. En consecuencia, *la dispersión de la distribución de medias muestrales es considerablemente menor que la dispersión de la población muestreada.*

El teorema del límite central se puede expresar de otro modo para aplicar a la **suma de las mediciones muestrales Σx_i**, que, cuando n se hace grande, también tiene una distribución aproximadamente normal con media $n\mu$ y desviación estándar $\sigma\sqrt{n}$.

La aportación importante del teorema del límite central está en la inferencia estadística. Muchos estimadores que se usan para hacer inferencias acerca de parámetros poblacionales son sumas o promedios de las mediciones muestrales. Cuando el tamaño muestral es lo suficientemente grande, se puede esperar que estos estimadores tengan distribuciones muestrales que sean aproximadamente normales. Entonces se puede usar la distribución normal para describir el comportamiento de estos estimadores en muestreo repetido y evaluar la probabilidad de observar ciertos resultados muestrales. Estas probabilidades se calculan usando la variable aleatoria normal estándar

$$z = \frac{\text{Estimador} - \text{Media}}{\text{Desviación estándar}}$$

Cuando usted vuelva a leer el teorema del límite central podrá observar que la aproximación es válida mientras el tamaño muestral n sea "grande", pero ¿qué tan grande es "grande"? Desafortunadamente, no hay una respuesta clara para esta pregunta. El valor apropiado de n depende de la forma de la población de la cual se muestrea, y cómo se desea usar la aproximación. No obstante, ayudan estas guías:

NECESITO SABER...

Cuándo el tamaño de la muestra es lo bastante grande para usar el teorema del límite central

- Si la población muestreada es **normal**, entonces la distribución muestral de $\bar{x}$ también será normal, sin importar cuál sea el tamaño de la muestra que se elija. Este resultado se puede demostrar en forma teórica, pero no debe ser demasiado difícil aceptarla sin prueba.
- Cuando la población muestreada es aproximadamente **simétrica**, la distribución muestral de $\bar{x}$ se hace también aproximadamente normal para valores relativamente pequeños de n. Recuerde la rapidez con la que la distribución "plana" del ejemplo de los dados tomó la forma de montículo ($n = 3$).
- Cuando la población muestreada está **sesgada**, el tamaño muestral n debe ser más grande, con n al menos 30 antes que la distribución muestral de $\bar{x}$ se haga aproximadamente normal.

Estas guías sugieren que, para muchas poblaciones, la distribución muestral de $\bar{x}$ será aproximadamente normal para tamaños muestrales moderados: una excepción a esta regla se presenta al muestrear una población binomial cuando p o $q = (1 - p)$ es muy pequeña. Cuando aparezcan aplicaciones específicas del teorema del límite central, daremos el tamaño muestral apropiado n.

2.4 DISTRIBUCIÓN MUESTRAL DE LA PROPORCIÓN

Hay numerosos ejemplos prácticos de la variable aleatoria x binomial. Una aplicación común se relaciona con la preferencia del consumidor o encuestas de opinión, donde usamos una muestra aleatoria de n personas, para estimar la proporción p de personas en la población que tienen una característica especificada. Si x de las personas muestreadas tienen esta característica, entonces la proporción muestral

$$\hat{p} = \frac{x}{n}$$

se puede usar para estimar la proporción poblacional p.[†]

TIP MI CONSEJO

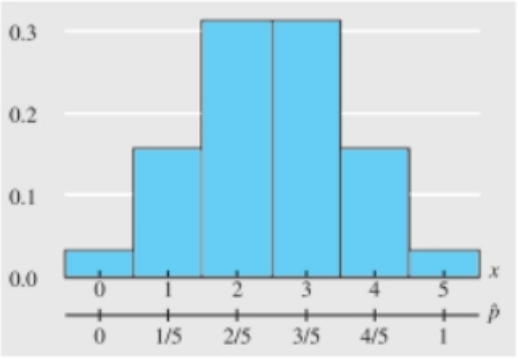
P: ¿Cómo saber si es o no binomial?

R: Observe si la medición tomada en una sola unidad experimental de la muestra es del tipo "éxito/fracaso". Si es así, probablemente es binomial

[†] Un "sombrero" puesto sobre el símbolo de un parámetro poblacional denota una estadística utilizada para estimar el parámetro poblacional. Por ejemplo, el símbolo $\hat{p}$ denota la proporción muestral.

La variable aleatoria x binomial tiene una distribución de probabilidad $p(x)$, con media np y desviación estándar $\sqrt{npq}$. Como $\hat{p}$ es simplemente el valor de x, expresado como proporción $\left(\hat{p} = \frac{x}{n}\right)$, la distribución muestral de $\hat{p}$ es idéntica a la distribución de probabilidad de x, excepto que tiene una nueva escala a lo largo del eje horizontal (figura 2.10).

FIGURA 2.10
Distribución muestral de la variable aleatoria x binomial y la proporción muestral $\hat{p}$



Debido a este cambio de escala, la media y la desviación estándar de $\hat{p}$ también tienen cambio de escala, de modo que la media de la distribución muestral de $\hat{p}$ es p, y su error estándar es

$$\text{SE}(\hat{p}) = \sqrt{\frac{pq}{n}} \qquad \text{donde } q = 1 - p$$

Por último, así como podemos aproximar la distribución de probabilidad de x con una distribución normal cuando el tamaño muestral n es grande, podemos hacer lo mismo con la distribución muestral de $\hat{p}$.

PROPIEDADES DE LA DISTRIBUCIÓN MUESTRAL DE LA PROPORCIÓN MUESTRAL, $\hat{p}$

- Si una muestra aleatoria de n observaciones se selecciona de una población binomial con parámetro p, entonces la distribución muestral de la proporción muestral

  $$\hat{p} = \frac{x}{n}$$

  tendrá una media

  $$p$$

  y una desviación estándar

  $$\text{SE}(\hat{p}) = \sqrt{\frac{pq}{n}} \qquad \text{donde } q = 1 - p$$

- Cuando el tamaño muestral n es grande, la distribución muestral de $\hat{p}$ puede ser aproximada por una distribución normal. La aproximación será adecuada si $np > 5$ y $nq > 5$.

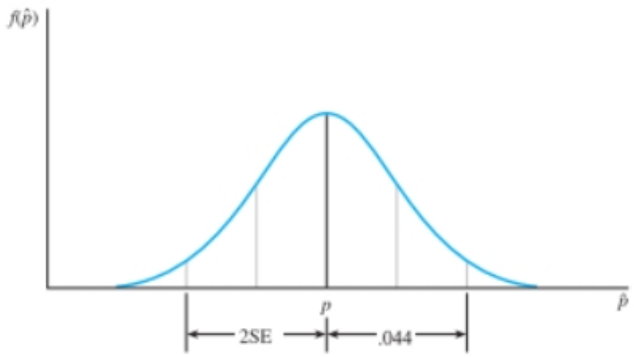
EJEMPLO 2.4

En una encuesta se preguntó a 500 madres y padres sobre la importancia del deporte para hijos e hijas. De los padres entrevistados 60% estuvo de acuerdo en que los géneros son iguales y deben tener iguales oportunidades de participar en deportes. Describa la distribución muestral de la proporción muestral $\hat{p}$ de padres que están de acuerdo en que los géneros son iguales y deben tener iguales oportunidades.

Solución Se puede suponer que los 500 padres representan una muestra aleatoria de los padres de todos los hijos e hijas en Estados Unidos y que la verdadera proporción de la población es igual a algún valor desconocido que se puede llamar p. La distribución muestral de $\hat{p}$ puede ser aproximada por una distribución muestral,[††] con media igual a p (véase la figura 2.11) y error estándar

$$\text{SE}(\hat{p}) = \sqrt{\frac{pq}{n}}$$

FIGURA 2.11
Distribución muestral para $\hat{p}$ basada en una muestra de $n = 500$ padres para el ejemplo 2.4



Se observa de la figura 2.11 que la distribución muestral de $\hat{p}$ está centrada sobre su media p. Aun cuando no se sabe el valor exacto de p (la proporción muestral $\hat{p} = .60$ puede ser mayor o menor que p), puede hallarse un valor aproximado para la desviación estándar de la distribución muestral usando la proporción muestral $\hat{p} = .60$ para aproximar el valor desconocido de p. Entonces,

$$\text{SE} = \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= \sqrt{\frac{(.60)(.40)}{500}} = .022$$

Por tanto, aproximadamente 95% de las veces, $\hat{p}$ caerá en no más de $2\text{SE} \approx .044$ del valor (desconocido) de p.

NECESITO SABER...

Cómo calcular probabilidades para la proporción muestral $\hat{p}$

1. Encuentre los valores necesarios de n y p.
2. Verifique si la aproximación normal a la distribución binomial es apropiada ($np > 5$ y $nq > 5$).
3. Escriba el evento de interés en términos de $\hat{p}$ y localice el área apropiada en la curva normal.
4. Convierta los valores necesarios de $\hat{p}$ en valores z usando
$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$
5. Use la tabla 3 del Apéndice para calcular la probabilidad.

[††] Si se verifican las condiciones que permiten la aproximación normal a la distribución de $\hat{p}$, se puede ver que $n = 500$ es adecuado para valores de p cercanos a .60 porque tanto $n\hat{p} = 300$ como $n\hat{q} = 200$ son mayores que 5.

EJEMPLO 2.5

Consulte el ejemplo 2.4. Suponga que la proporción p de padres en la población es en realidad igual a .55. ¿Cuál es la probabilidad de observar una proporción muestral igual de grande o mayor que el valor observado $\hat{p} = .60$?

Solución La figura 2.12 muestra la distribución muestral de $\hat{p}$ cuando $p = .55$, con el valor observado $\hat{p} = .60$ ubicado sobre el eje horizontal. La probabilidad de observar una proporción muestral $\hat{p}$ igual a o mayor que .60 es aproximada por el área sombreada en la cola superior de esta distribución normal con

$$p = .55$$

y

$$\text{SE} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.55)(.45)}{500}} = .0222$$

FIGURA 2.12
Distribución muestral de $\hat{p}$ para $n = 500$ y $p = .55$ para el ejemplo 2.5

Para hallar el área sombreada, primero calcule el valor z correspondiente a $\hat{p} = .60$:

$$z = \frac{\hat{p} - p}{\sqrt{pq/n}} = \frac{.60 - .55}{.0222} = 2.25$$

Usando la tabla 3 del Apéndice, encontramos

$$P(\hat{p} > .60) \approx P(z > 2.25) = 1 - .9878 = .0122$$

Esto es, si seleccionáramos una muestra aleatoria de $n = 500$ observaciones de una población con proporción p igual a .55, la probabilidad de que la proporción muestral $\hat{p}$ sería tan grande o mayor que .60 es de sólo .0122.

La corrección equivalente aquí es $\pm(.5/n)$. Por ejemplo, para $\hat{p} = .60$ el valor de z con la corrección es

$$z_1 = \frac{(.60 - .001) - .55}{\sqrt{\frac{(.55)(.45)}{500}}} = 2.20$$

con $P(\hat{p} > .60) \approx .0139$. Con una precisión de dos lugares decimales, este valor concuerda con el resultado anterior. Cuando n es grande, el efecto de usar la corrección es por lo general insignificante. Usted debe resolver problemas de este capítulo y los siguientes *sin* el factor de corrección a menos que específicamente se le solicite.

EJERCICIOS

TÉCNICAS BÁSICAS

2.1 Se seleccionaron muestras aleatorias de tamaño n de poblaciones con las medias y varianzas dadas aquí. Encuentre la media y desviación estándar de la distribución muestral de la media muestral en cada caso:

a. $n = 36, \mu = 10, \sigma^2 = 9$

b. $n = 100, \mu = 5, \sigma^2 = 4$

c. $n = 8, \mu = 120, \sigma^2 = 1$

DATOS EX0717

2.2 Una población está formada por $N = 5$ números: 1, 3, 5, 6 y 7. Se puede demostrar que la media y desviación estándar para esta población son $\mu = 4.4$ y $\sigma = 2.15$, respectivamente.

a. Construya un histograma de probabilidad para esta población.

b. Use la tabla de números aleatorios, tabla 10 del Apéndice, para seleccionar una muestra aleatoria de tamaño $n = 10$ con remplazo de la población. Calcule la media muestral, $\bar{x}$. Repita este procedimiento, calculando la media muestral $\bar{x}$ para su segunda muestra. (SUGERENCIA: Asigne los dígitos aleatorios 0 y 1 a la medición $x = 1$; asigne dígitos 2 y 3 a la medición $x = 3$ y así sucesivamente.)

c. Para simular la distribución muestral de $\bar{x}$, hemos seleccionado 50 muestras más de tamaño $n = 10$ con restitución y hemos calculado las correspondientes medias muestrales. Construya un histograma de frecuencia relativa para estos 50 valores de $\bar{x}$. ¿Cuál es la forma de esta distribución?

4.8	4.2	4.2	4.5	4.3	4.3	5.0	4.0	3.3	4.7
3.0	5.9	5.7	4.2	4.4	4.8	5.0	5.1	4.8	4.2
4.6	4.1	3.4	4.9	4.1	4.0	3.7	4.3	4.3	4.5
5.0	4.6	4.1	5.1	3.4	5.9	5.0	4.3	4.5	3.9
4.4	4.2	4.2	5.2	5.4	4.8	3.6	5.0	4.5	4.9

2.3 Una muestra aleatoria de n observaciones se selecciona de una población con desviación estándar $\sigma = 1$. Calcule el error estándar de la media (SE) para estos valores de n:

a. $n = 1$ **b.** $n = 2$ **c.** $n = 4$

d. $n = 9$ **e.** $n = 16$ **f.** $n = 25$

g. $n = 100$

2.4 Se selecciona una muestra aleatoria de tamaño $n = 49$ de una población con media $\mu = 53$ y desviación estándar $\sigma = 21$.

a. ¿Cuál será la forma aproximada de la distribución muestral de $\bar{x}$?

b. ¿Cuáles serán la media y la desviación estándar de la distribución muestral de $\bar{x}$?

2.5 Consulte el ejercicio 2.4. Encuentre la probabilidad de que la media muestral sea mayor que 55.

2.6 Suponga que se selecciona una muestra aleatoria de $n = 25$ observaciones de una población que está distribuida normalmente, con media igual a 106 y desviación estándar igual a 12.

a. Dé la media y desviación estándar de la distribución muestral de la media muestral $\bar{x}$.

b. Encuentre la probabilidad de que $\bar{x}$ exceda de 110.

c. Encuentre la probabilidad de que la media muestral se desvíe de la media poblacional $\mu = 106$ en no más de 4.

2.7 Se seleccionaron muestras aleatorias de tamaño n de poblaciones binomiales con parámetros poblacionales p dados aquí. Encuentre la media y la desviación estándar de la distribución muestral de la proporción muestral $\hat{p}$ en cada caso:

a. $n = 100, p = .3$

b. $n = 400, p = .1$

c. $n = 250, p = .6$

2.8 Calcule SE $(\hat{p})$ para $n = 100$ y estos valores de p:

a. $p = .01$ **b.** $p = .10$ **c.** $p = .30$

d. $p = .50$ **e.** $p = .70$ **f.** $p = .90$

g. $p = .99$

2.9 **a.** ¿Es apropiada la aproximación normal a la distribución muestral de $\hat{p}$ cuando $n = 400$ y $p = .8$?

b. Use los resultados de la parte **a** para hallar la probabilidad de que $\hat{p}$ es mayor que .83.

c. Use los resultados de la parte **a** para hallar la probabilidad de que $\hat{p}$ está entre .76 y .84.

APLICACIONES

2.10 **Salarios de profesorado** Suponga que el profesorado de una universidad, con el rango de profesor en instituciones públicas que imparten carreras de dos años, ganan un promedio de 71 802 dólares por año[1] con una desviación estándar de 4 000 dólares. En un intento por verificar este nivel de salario se seleccionó una muestra aleatoria de 60 profesores de entre una base de datos del personal para todas las instituciones de dos años en Estados Unidos.

a. Describa la distribución muestral de la media muestral $\bar{x}$.

b. ¿Dentro de qué límites se esperaría que esté el promedio muestral, con probabilidad .95?

c. Calcule la probabilidad de que la media muestral $\bar{x}$ sea mayor que 73 000 dólares.

d. Si su muestra aleatoria en realidad produjo una media muestral de 73 000 dólares, ¿consideraría usted que esto es poco común? ¿Qué conclusión obtendría?

2.11 **Bacterias en el agua** Use el teorema del límite central para explicar por qué una variable aleatoria de Poisson, por ejemplo, el número de un tipo particular de bacterias en un pie cúbico de agua, tiene una distribución que puede ser aproximada por una distribución normal cuando la media μ es grande. (SUGERENCIA: Un pie cúbico de agua contiene 1 728 pulgadas cúbicas de líquido.)

2.12 **Venta de bocadillos** El total de ventas diarias x en la sección de bocadillos de un mercado local es la suma de las ventas generadas por un número fijo de clientes que hacen compras en un día determinado.

a. ¿Qué clase de distribución de probabilidad se espera que tenga el total de ventas diarias? Explique.

b. Para este mercado particular, el promedio de venta por cliente en la sección de bocadillos es \$8.50 con σ = \$2.50. Si 30 clientes hacen compras de bocadillos en un día determinado, dé la media y la desviación estándar de la distribución de probabilidad del total de ventas diarias, x.

2.13 M&M's Un publicista afirma que el porcentaje promedio de dulces M&M's cafés de un paquete de chocolates M&M's de leche es 13%. Suponga que se selecciona al azar un paquete de chocolates M&M's de leche que contiene 55 dulces y se determina la proporción de dulces M&M's cafés del paquete.

a. ¿Cuál es la distribución aproximada de la proporción muestral de dulces M&M's cafés en un paquete que contiene 55 dulces?

b. ¿Cuál es la probabilidad de que el porcentaje muestral de dulces cafés sea menor que 20%?

c. ¿Cuál es la probabilidad de que el porcentaje muestral exceda de 35%?

d. ¿Dentro de qué intervalo espera usted que la proporción muestral se encuentre alrededor de 95% de las veces?

2.14 Fido en el auto Parece que conducir con una mascota en el auto es la tercera peor distracción, detrás de hablar por teléfono y enviar mensajes de texto.

Fuente: USA Today, 19 de agosto de 2010, p. 8A

Según un estudio de la American Automobile Association, 80% de los conductores admiten conducir con una mascota en el auto, y de éstos, 20% permite que su perro se siente en su regazo.[2] Suponga que selecciona al azar una muestra de 100 conductores que han admitido conducir con una mascota en su auto.

a. ¿Cuál es la probabilidad de que 25% o más de los conductores permitan que sus perros se sienten en sus regazos?

b. ¿Cuál es la probabilidad de que 10% o menos de los conductores permitan que sus perros se sienten en sus regazos?

c. ¿Sería inusual encontrar que 35% de los conductores permitan que sus perros se sienten en sus regazos?

2.15 ¡Ah, las nueces! ¿Es usted un "purista" del chocolate o le gusta que tenga otros ingredientes? *American Demographics* informa que casi 75% de los consumidores gustan de ingredientes tradicionales como nueces o caramelos en su chocolate. Son menos entusiastas hacia el gusto de la menta o el café, que dan sabores más distintivos.[3] Se selecciona una muestra aleatoria de 200 consumidores y se registra el número de quienes gustan de las nueces o caramelo en su chocolate.

a. ¿Cuál es la distribución muestral aproximada para la proporción muestral $\hat{p}$? ¿Cuáles son la media y la desviación estándar para esta distribución?

b. ¿Cuál es la probabilidad de que el porcentaje muestral sea mayor que 80%?

c. ¿Dentro de qué límites se esperaría que la proporción muestral se encuentre alrededor de 95% de las veces?

Fuentes

1. "Average Salary of Full Time Instructional Faculty on 9-Month Contracts in Degree-granting Institutions," *Digest of Educational Statistics*, http://nces.ed.gov/programs/digest/d09/tables/dt09_258.asp, 5 de julio de 2010.
2. Meghan Baker, "We're Driven to Distraction When Fido is Co-Pilot, Study Finds," www.foxnews.com/us/2010/08/19/driven-distraction-pets-car-just-dangerour-texting/, 20 de agosto de 2010.
3. Adam Fernandez, "Nuts About You," *American Demographics* 26, no. 1 (febrero de 2004):14.

3

Inferencia estadística con muestras grandes

© Associated Press

OBJETIVOS GENERALES

Usted ya conoce las distribuciones de probabilidad de variables aleatorias y las distribuciones muestrales de varias estadísticas que, para tamaños muestrales grandes, pueden ser aproximadas por una distribución normal de acuerdo con el teorema del límite central. Este capítulo presenta un método para estimar parámetros poblacionales e ilustra el concepto con ejemplos prácticos. El teorema del límite central y las distribuciones muestrales presentadas en el capítulo 2 desempeñan un papel clave para evaluar la confiabilidad de estimaciones.

ÍNDICE DEL CAPÍTULO

- Estimación de intervalos
- Estimación de la diferencia entre dos medias poblacionales
- Estimación de la diferencia entre dos proporciones binomiales
- Estimación puntual para una media o proporción poblacional
- Intervalos de confianza de muestra grande para una media o proporción poblacional
- Selección del mejor estimador puntual
- Selección del tamaño muestral

NECESITO SABER...

Cómo estimar una media o proporción poblacional
Cómo seleccionar el tamaño muestral

INFERENCIA ESTADÍSTICA, TIPOS DE ESTIMADORES, EVALUACIÓN DE LA BONDAD DE UN ESTIMADOR PUNTUAL Y DE UN ESTIMADOR POR INTERVALO

3.1

El primer capítulo de este libro ha proporcionado el material que necesitará para entender la inferencia estadística y cómo se aplica en situaciones prácticas. En él se proporcionó el vínculo entre la probabilidad y la inferencia estadística. Numerosas estadísticas son sumas o promedios calculados de mediciones muestrales. El teorema del límite central dice que, incluso si las poblaciones no son normales, las distribuciones de esas *estadísticas* serán aproximadamente normales cuando el tamaño muestral n es grande. Estas estadísticas son las herramientas que se usarán para las *estadísticas inferenciales*, es decir, hacer inferencias acerca de una población usando información contenida en una muestra.

Inferencia estadística

La inferencia, específicamente la toma de decisiones y la predicción, tiene siglos de antigüedad y desempeña un papel muy importante en la vida de casi todas las personas. Veamos a continuación algunas aplicaciones:

- El gobierno necesita predecir las tasas de interés a corto y largo plazos.
- Un corredor financiero desea pronosticar el comportamiento del mercado de acciones.
- Un metalurgista desea determinar si un nuevo tipo de acero es más resistente a altas temperaturas que el actual.
- Una consumidora desea estimar el precio de venta de su casa antes de ponerla en el mercado.

TIP MI CONSEJO

Parámetro ⇔ población
Estadística ⇔ muestra

Hay muchas formas de tomar estas decisiones o predicciones, algunas son subjetivas y otras son de naturaleza más objetiva. ¿Qué tan buenas serán las predicciones o decisiones? Aun cuando usted piense que su propia capacidad de tomar decisiones es muy buena, la experiencia sugiere que éste puede no ser el caso. Es la función del estadístico matemático dar métodos para hacer inferencia estadística que son mejores y más confiables que únicamente las suposiciones subjetivas.

La inferencia estadística se ocupa de tomar decisiones o hacer predicciones acerca de **parámetros**, es decir, las medidas numéricas descriptivas que caracterizan a una población. Los parámetros que encontramos en capítulos anteriores son la media poblacional μ, la desviación poblacional estándar σ y la proporción binomial p. En inferencia estadística, un problema práctico se expone de otra forma en el marco de una población con un parámetro específico de interés. Por ejemplo, el metalurgista podría medir el *promedio* de coeficientes de expansión de ambos tipos de acero y luego comparar sus valores.

Los métodos para hacer inferencias acerca de parámetros poblacionales caen en una de dos categorías:

- **Estimación:** Estimar o predecir el valor del parámetro
- **Prueba de hipótesis:** Tomar una decisión acerca del valor de un parámetro, con base en alguna idea preconcebida acerca de cuál podría ser su valor

EJEMPLO 3.1

Los circuitos en computadoras y otros equipos electrónicos están formados por una o más tarjetas de circuito impreso (PCB) y es frecuente que las computadoras sean reparadas con sólo cambiar una o más tarjetas defectuosas. En un intento por hallar el ajuste apropiado de un proceso de chapa aplicado a uno de los lados de una PCB, un supervisor de producción podría *estimar* el grosor promedio de chapa de cobre en las PCB usando muestras de varios días de

operación. Como no sabe del grosor promedio μ antes de observar el proceso de producción, su problema es de *estimación*.

EJEMPLO 3.2

El supervisor del ejemplo 3.1 recibe instrucciones del propietario de la planta de que el grosor de la chapa de cobre no debe ser menor que .001 de pulgada, para que el proceso esté en control. Para decidir si el proceso está o no en control, el supervisor debe formular una prueba. Podría *hacer una hipótesis* de que el proceso está en control, es decir, suponer que el grosor promedio de la chapa de cobre es .001 o mayor, y usar muestras de varios días de operación para decidir si es o no correcta su hipótesis. El método de la toma de decisión del supervisor se denomina *prueba de hipótesis.*

¿Cuál método de inferencia debe usarse? Esto es, ¿debe estimarse el parámetro o probar una hipótesis respecto a su valor? La respuesta está dictada por la pregunta práctica planteada y a veces es determinada por preferencias personales. Como la estimación y las pruebas de hipótesis se usan con frecuencia en literatura científica, incluimos ambos métodos en este capítulo.

Un problema estadístico, que comprende planeación, análisis y formulación de inferencias, está incompleto sin una medida de la **bondad de la inferencia**. Esto es, ¿qué tan preciso o confiable es el método empleado? Si una corredora financiera predice que el precio de una acción será de 80 dólares el próximo lunes, ¿estaría usted dispuesto a comprar o vender su acción sin saber qué tan confiable es la predicción? ¿La predicción estará a no más de 1, 2 o 10 dólares del precio real el próximo lunes? Los procedimientos estadísticos son importantes porque dan dos tipos de información:

- Métodos para hacer la inferencia
- Una medida numérica de la bondad o confiabilidad de la inferencia

Tipos de estimadores

Para estimar el valor de un parámetro poblacional puede usar información de la muestra en la forma de un **estimador**. Los estimadores se calculan usando información de las observaciones muestrales y, en consecuencia, por definición son también *estadísticas*.

Definición Un **estimador** es una regla, generalmente expresada como fórmula, que nos dice cómo calcular una estimación basada en información de la muestra.

Los estimadores se usan en dos formas diferentes:

- **Estimación puntual:** Con base en datos muestrales, se calcula un solo número para estimar el parámetro poblacional. La regla o fórmula que describe este cálculo se denomina **estimador puntual** y el número resultante recibe el nombre de **estimación puntual**.
- **Estimación de intervalo:** Con base en datos muestrales, dos números se calculan para formar un intervalo dentro del cual se espera esté el parámetro. La regla o fórmula que describe este cálculo se denomina **estimador de intervalo** y el par de números resultantes se llama **estimación de intervalo** o **intervalo de confianza**.

EJEMPLO 3.3

Un veterinario desea estimar el aumento mensual promedio en el peso de cachorros de raza golden retriever, de cuatro meses de edad, que han sido puestos a dieta de carne de cordero y arroz. La *población* está formada por los aumentos mensuales en el peso de todos los cacho-

rros de raza golden retriever a los que se da esta dieta particular. El veterinario desea estimar el parámetro desconocido μ, el aumento mensual promedio en el peso para esta población *hipotética*.

Un posible *estimador* basado en datos muestrales es la media muestral, $\bar{x} = \Sigma x_i/n$. Podría usarse en la forma de un solo número o *estimación puntual*, por ejemplo, 3.8 libras, o usarse una *estimación de intervalo* y estimar que el aumento promedio en el peso será entre 2.7 y 4.9 libras.

Los procedimientos de estimación tanto puntuales como de intervalo usan información dada por la distribución muestral del estimador específico que se haya elegido para usarse. Empezaremos por exponer la *estimación puntual* y su uso para estimar medias poblacionales y proporciones.

3.2 ESTIMACIÓN PUNTUAL DE LA MEDIA DE UNA POBLACIÓN, ESTIMACIÓN POR INTERVALO DE LA MEDIA DE UNA POBLACIÓN

TIP MI CONSEJO

Parámetro = diana del blanco

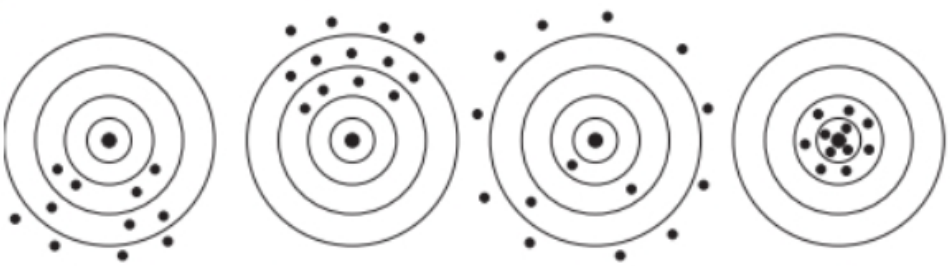
Estimador = bala o flecha

En una situación práctica, existen varias estadísticas que podrían usarse como estimadores puntuales para un parámetro poblacional. Para determinar cuál de las opciones es mejor, usted necesita saber cómo se comporta el estimador en muestreo repetido, descrito por su *distribución muestral*.

Como una analogía, considere disparar con un revólver hacia un blanco. El parámetro de interés es la diana a la cual se disparan balas. Cada bala representa una sola estimación muestral, disparada por el revólver, que representa el estimador.

Suponga que un amigo dispara una sola bala y acierta en la diana. ¿Se puede concluir que él es un excelente tirador? ¿Se pondría usted de pie junto al blanco cuando él dispare una segunda bala? Es probable que no, porque no tiene una medida de lo bien que él dispare en intentos repetidos. ¿Siempre acierta en el blanco o sus tiros son demasiado altos o demasiado bajos en forma consistente? ¿Sus tiros se agrupan alrededor del blanco o fallan en forma consistente en acertar en el blanco por un amplio margen? La figura 3.1 muestra varias configuraciones del blanco. ¿Cuál blanco elegiría usted como perteneciente al mejor tiro?

FIGURA 3.1
¿Cuál tirador es el mejor?

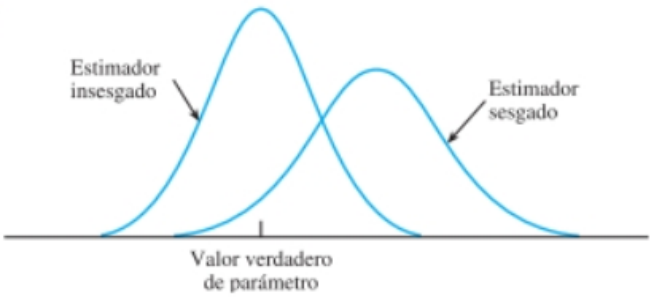


Las distribuciones muestrales dan información que se puede usar para seleccionar el **mejor estimador**. ¿Qué características serían valiosas? Primero, la **distribución muestral del estimador puntual debe estar centrada sobre el valor verdadero del parámetro que se estimará**. Esto es, el estimador no debe subestimar o sobreestimar de manera constante al parámetro de interés. Un estimador como éste se dice que es **insesgado**.

Definición Se dice que un estimador de un parámetro es **insesgado** si la media de su distribución es igual al valor verdadero del parámetro. De otro modo, se dice que el estimado está **sesgado**.

Las distribuciones muestrales para un estimador insesgado y un estimador sesgado se presentan en la figura 3.2. La distribución muestral para el estimador sesgado se encuentra hacia la derecha del valor verdadero del parámetro. Este estimador sesgado es más probable que uno insesgado para sobreestimar el valor del parámetro.

FIGURA 3.2
Distribuciones para estimadores sesgados e insesgados



La segunda característica deseable de un estimador es que **la dispersión (medida por la varianza) de la distribución muestral debe ser tan pequeña como sea posible**. Esto asegura que, con una alta probabilidad, una estimación individual quedará cerca del valor verdadero del parámetro. Las distribuciones muestrales para dos estimadores insesgados, una con una varianza pequeña[†] y la otra con una varianza más grande, se presentan en la figura 3.3.

FIGURA 3.3
Comparación de variabilidad de un estimador

Por supuesto que sería preferible el estimador con la varianza más pequeña, porque las estimaciones tienden a estar más cerca del valor verdadero del parámetro que en la distribución con la varianza más grande.

En situaciones muestrales prácticas es posible saber que la distribución muestral de un estimador está centrada alrededor del parámetro que se trate de estimar, pero todo lo que se tiene es la estimación calculada de las n mediciones contenidas en la muestra. ¿A qué distancia del

[†] En general, los estadísticos usan el término *varianza de un estimador* cuando en realidad es la varianza de la distribución muestral del estimador. Esta expresión contraída se usa casi universalmente.

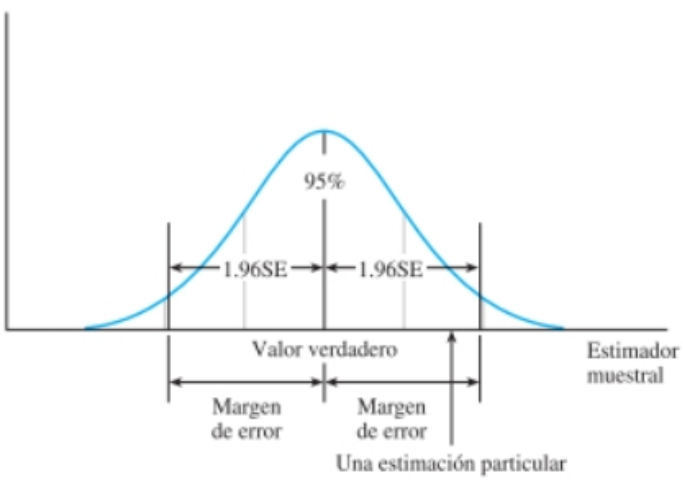
valor verdadero del parámetro estará esta estimación? ¿Qué tan cercana está la bala del tirador de la diana o blanco? La distancia entre la estimación y el valor verdadero del parámetro se denomina **error de estimación**.

Definición La distancia entre una estimación y el parámetro estimado recibe el nombre de **error de estimación**.

En este capítulo, usted puede suponer que los tamaños muestrales son siempre grandes y, por tanto, que los estimadores *insesgados* que estudiará tienen distribuciones muestrales que pueden ser aproximadas por una distribución normal (por el teorema del límite central). Recuerde que, para cualquier estimador puntual con una distribución normal, la regla empírica dice que aproximadamente 95% de todas las estimaciones puntuales estarán a no más de dos (o más exactamente, 1.96) desviaciones estándar de la media de esa distribución.

Para estimadores *insesgados*, esto implica que la diferencia entre el estimador puntual y el valor verdadero del parámetro será menor que 1.96 desviaciones estándar o 1.96 errores estándar (SE). Esta cantidad, llamada el 95% de **margen de error** (o simplemente "**margen de error**"), da un límite superior práctico para el error de estimación (véase la figura 3.4). Es posible que el error de estimación exceda este margen de error, pero eso es muy poco probable.

FIGURA 3.4
Distribución muestral de un estimador insesgado



TIP MI CONSEJO
95% de margen de error = 1.96 × error estándar

ESTIMACIÓN PUNTUAL DEL PARÁMETRO DE UNA POBLACIÓN

- Estimador puntual: estadística calculada usando mediciones muestrales
- 95% de margen de error: 1.96 × error estándar del estimador

Se puede demostrar que estos dos estimadores puntuales tienen la *mínima variabilidad* de todos los estimadores insesgados y, por lo tanto, son los *mejores estimadores* que se pueden hallar en cada situación.

La variabilidad del estimador se mide usando este error estándar. No obstante, usted habrá observado que el error estándar suele depender de parámetros desconocidos, por ejemplo σ o p. Estos parámetros deben estimarse usando estadísticas muestrales como s y $\hat{p}$. Aun cuando no es exactamente correcto, por lo general los experimentadores se refieren al error estándar estimado como *el error estándar*.

NECESITO SABER...

Cómo estimar una media o proporción poblacional

- Para estimar la media poblacional μ para una población cuantitativa, el estimador puntual $\bar{x}$ es *insesgado* con el error estándar estimado como

$$\text{SE} = \frac{s}{\sqrt{n}}^{\dagger}$$

El 95% de margen de error cuando $n \geq 30$ se estima como

$$\pm 1.96\left(\frac{s}{\sqrt{n}}\right)$$

- Para estimar la proporción poblacional p para una población binomial, el estimador puntual $\hat{p} = x/n$ es insesgado, con un error estándar estimado como

$$\text{SE} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

El 95% de margen de error se estima como

$$\pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Suposiciones: $n\hat{p} > 5$ y $n\hat{q} > 5$.

EJEMPLO 3.4

Un ambientalista realiza un estudio del oso polar, especie que se encuentra en el océano Ártico y sus alrededores. Su zona de distribución está limitada por la existencia de hielo en el mar, que usa como plataforma para cazar focas, principal sostén de su dieta. La destrucción de su hábitat en el hielo del Ártico, que se ha atribuido al calentamiento global, amenaza la supervivencia de los osos como especie; puede extinguirse antes de un siglo.[†] Una muestra aleatoria de $n = 50$ osos polares produjo un peso promedio de 980 libras con una desviación estándar de 105 libras. Use esta información para estimar el peso promedio de todos los osos polares del Ártico.

Solución La variable aleatoria medida es el peso, una variable aleatoria cuantitativa mejor descrita por su media μ. La estimación puntual de μ, el peso promedio de todos los osos polares del Ártico, es $\bar{x} = 980$ libras. El margen de error se estima como

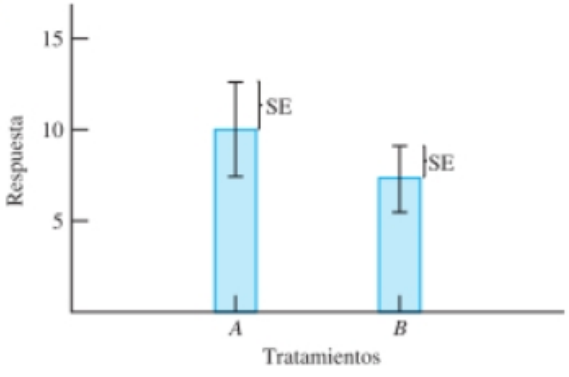
$$1.96\,\text{SE} = 1.96\left(\frac{s}{\sqrt{n}}\right) = 1.96\left(\frac{105}{\sqrt{50}}\right) = 29.10 \approx 29 \text{ libras}$$

Se puede tener confianza en que la estimación muestral de 980 libras está a no más de ± 29 libras de la media poblacional.

[†] Cuando se muestrea a partir de una distribución normal, el estadístico $(\bar{x} - \mu)/(s/\sqrt{n})$ tiene una distribución t, que se estudiará en el capítulo 4. Cuando la muestra es *grande*, este estadístico se encuentra distribuido normalmente en forma aproximada si la población muestreada es normal o no normal.

Al informar los resultados de una investigación es frecuente que los investigadores agreguen ya sea la desviación muestral estándar s (a veces llamada SD) o el error estándar $s/\sqrt{n}$ (por lo general llamado SE o SEM) a las estimaciones de medias poblacionales. Siempre se debe buscar una explicación en el texto del informe que diga si el investigador está informando $\bar{x} \pm$ SD o $\bar{x} \pm$ SE. Además, las medias muestrales y desviaciones estándar o errores estándar se presentan a menudo como "barras de error" usando el formato gráfico que se ilustra en la figura 3.5.

FIGURA 3.5
Gráfica de medias de tratamiento y sus errores estándar



EJEMPLO 3.5

Además del peso promedio del oso polar del Ártico, el ambientalista del ejemplo 3.4 también está interesado en las opiniones de adultos sobre el tema del calentamiento global. En particular, desea estimar la proporción de personas que piensan que el calentamiento global es un problema muy serio. En una muestra aleatoria de $n = 100$ adultos, 73% indicó que el calentamiento global es un problema muy serio. Estime la verdadera proporción de población de adultos que piensan que el calentamiento global es un problema muy serio y encuentre el margen de error para la estimación.

Solución El parámetro de interés es ahora p, la proporción de personas en la población que piensan que el calentamiento global es un problema muy serio. El mejor estimador de p es la proporción muestral $\hat{p}$, que para esta muestra es $\hat{p} = .73$. Para hallar el margen de error, usted puede aproximar el valor de p con su estimación $\hat{p} = .73$:

$$1.96 \text{ SE} = 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{.73(.27)}{100}} = .09$$

Con este margen de error, se puede estar bastante seguro de que la estimación de .73 está dentro de $\pm .09$ del valor verdadero de p. En consecuencia, se puede concluir que el valor verdadero de p podría ser de sólo .64 o de hasta .82. Este margen de error es bastante grande cuando se compara con la estimación misma y refleja el hecho de que se requiere de muestras grandes para alcanzar un margen de error pequeño cuando se estime p.

TABLA 3.1 **Algunos valores calculados de $\sqrt{pq}$**

p	pq	$\sqrt{pq}$	p	pq	$\sqrt{pq}$
.1	.09	.30	.6	.24	.49
.2	.16	.40	.7	.21	.46
.3	.21	.46	.8	.16	.40
.4	.24	.49	.9	.09	.30
.5	.25	.50			

La tabla 3.1 muestra la forma en que el numerador del error estándar de $\hat{p}$ cambia para diversos valores de p. Observe que, para casi todos los valores de p, en especial cuando p está entre .3 y .7, hay muy poco cambio en $\sqrt{pq}$, el numerador del SE, que alcanza su máximo valor cuando $p = .5$. Esto significa que el margen de error usando el estimador $\hat{p}$ también será máximo cuando $p = .5$. Cuando estiman p, algunos entrevistadores rutinariamente usan el margen máximo de error, que a veces recibe el nombre de **error muestral**, en cuyo caso calculan

$$1.96\,\text{SE} = 1.96\sqrt{\frac{.5(.5)}{n}} \quad \text{o a veces} \quad 2\,\text{SE} = 2\sqrt{\frac{.5(.5)}{n}}$$

Las encuestas Gallup, Harris y Roper generalmente usan tamaños muestrales de alrededor de 1000, de modo que su margen de error es

$$1.96\sqrt{\frac{.5(.5)}{1000}} = .031 \quad \text{o sea alrededor de 3\%}$$

En este caso, se dice que la estimación está dentro de ± 3 puntos porcentuales de la verdadera proporción de población.

Estimación por intervalo

Un *estimador de intervalo* es una regla para calcular dos números, por ejemplo a y b, para crear un intervalo del que usted esté bastante seguro que contiene el parámetro de interés. El concepto de "bastante seguro" significa "con gran probabilidad". Medimos esta probabilidad usando el **coeficiente de confianza**, designado por $\mathbf{1 - \alpha}$.

Definición La probabilidad de que un intervalo de confianza contenga el parámetro estimado se denomina **coeficiente de confianza**.

TIP MI CONSEJO

Cómo lazar:

Parámetro = poste de cerca

Estimación de intervalo = lazo

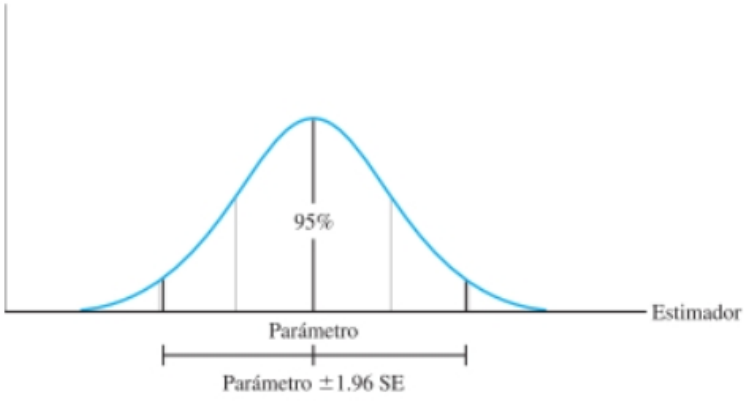
Por ejemplo, es frecuente que los experimentadores construyan intervalos de confianza de 95%, lo cual significa que el coeficiente de confianza, o la probabilidad de que el intervalo contenga el parámetro estimado, es .95. Puede aumentar o reducir la cantidad de certeza si se cambia el coeficiente de confianza. Algunos valores que por lo general usan experimentadores son .90, .95, .98 y .99.

Considere una analogía, esta vez lanzar un lazo a un poste de una cerca. El poste de la cerca representa el parámetro que se desea estimar y el lazo formado por la cuerda representa el intervalo de confianza. Cada vez que se lance la cuerda, se espera lazar al poste de la cerca; no obstante, a veces falla el lazo. En la misma forma, cada vez que se saque una muestra y construya un intervalo de confianza para un parámetro, usted espera incluir el parámetro en su intervalo, pero, al igual que el lazo, a veces falla. Su "porcentaje de éxito", es decir la proporción de intervalos que "lazan al poste" en muestreo repetido, es el coeficiente de confianza.

Construcción de un intervalo de confianza

Cuando la distribución muestral de un estimador puntual es aproximadamente normal, se puede construir un estimador de intervalo o **intervalo de confianza** mediante el siguiente razonamiento. Para mayor sencillez, suponga que el coeficiente de confianza es .95 y consulte la figura 3.6.

FIGURA 3.6
Parámetro ±1.96 SE



- Sabemos que, de todos los valores posibles del estimador que podríamos seleccionar, 95% de ellos estarán en el intervalo

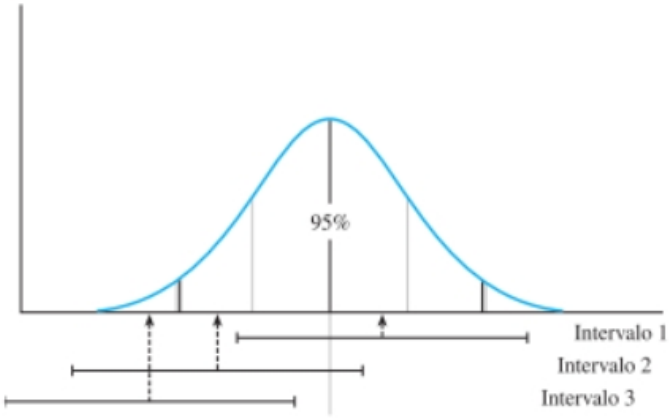
  Parámetro ±1.96 SE

  que se muestra en la figura 3.6.
- Como el valor del parámetro es desconocido, considere construir el intervalo

  Estimador ±1.96 SE

  que tiene el mismo ancho que el primer intervalo, pero tiene un centro variable.
- ¿Con qué frecuencia funcionará este intervalo en forma correcta y encerrará el parámetro de interés? Consulte la figura 3.7.

FIGURA 3.7
Algunos intervalos de confianza de 95%

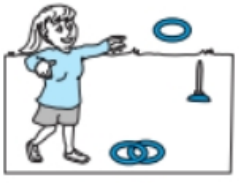


TIP MI CONSEJO

Al igual que en un juego de lanzar un anillo:

Parámetro = estaquilla

Estimación de intervalo = anillo



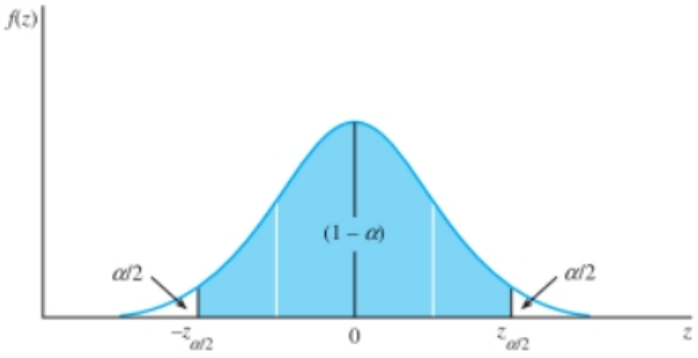
Los primeros dos intervalos funcionan correctamente, es decir, el parámetro (marcado con una línea gris claro) está contenido dentro de ambos intervalos. El tercer intervalo no funciona, porque no encierra al parámetro. Esto ocurrió porque el valor del estimador del centro del intervalo estaba demasiado lejos del parámetro. Por fortuna, valores del estimador sólo caen a esa distancia 5% de las veces y ¡nuestro procedimiento funcionará en forma correcta 95% del tiempo!

Si lo desea puede cambiar el *coeficiente de confianza* de $(1 - \alpha) = .95$ a otro nivel de confianza $(1 - \alpha)$. Para lograr esto, es necesario cambiar el valor $z = 1.96$, que localiza un área de .95 en el centro de la curva normal estándar, a un valor de z que localice el área $(1 - \alpha)$ en el centro de la curva, como se muestra en la figura 3.8.

Como el área total bajo la curva es 1, el área restante en las dos colas es α y cada cola contiene un área $\alpha/2$. El valor de z que tiene "área de cola" $\alpha/2$ a su derecha se denomina $z_{\alpha/2}$, y el área entre $-z_{\alpha/2}$ y $z_{\alpha/2}$ es el coeficiente de confianza $(1 - \alpha)$. Valores de $z_{\alpha/2}$, que por lo general son

utilizados por experimentadores, le serán familiares cuando empiece a construir intervalos de confianza para diferentes situaciones prácticas. Algunos de estos valores se dan en la tabla 3.2.

FIGURA 3.8
Ubicación de $z_{\alpha/2}$



INTERVALO DE CONFIANZA DE MUESTRA GRANDE (1 – α)100%

(Estimador puntual) $\pm$ $z_{\alpha/2}$ $\times$ (error estándar del estimador)

donde $z_{\alpha/2}$ es el valor z con un área $\alpha/2$ en la cola derecha de una distribución normal estándar. Esta fórmula genera dos valores; el **límite inferior de confianza (LCL)** y el **límite superior de confianza (UCL)**.

TABLA 3.2 **Valores de *z* que comúnmente se usan para intervalos de confianza**

Cociente de confianza $(1 - \alpha)$	α	$\alpha/2$	$z_{\alpha/2}$
.90	.10	.05	1.645
.95	.05	.025	1.96
.98	.02	.01	2.33
.99	.01	.005	2.58

Intervalo de confianza de muestra grande para una media poblacional μ

Es muy frecuente que problemas prácticos lleven a la estimación de μ, la media de una población de mediciones cuantitativas. He aquí algunos ejemplos:

- El promedio de calificaciones de estudiantes universitarios en una universidad particular
- El promedio de resistencia de un nuevo tipo de acero
- El número promedio de fallecimientos por categoría de edad
- El promedio de demanda para un nuevo producto de cosmético

Cuando el tamaño muestral n es grande, la media muestral $\bar{x}$ es el mejor estimador puntual para la media poblacional μ. Como su distribución muestral es aproximadamente normal, puede usarse para construir un intervalo de confianza de acuerdo con el método general dado ya antes.

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE (1 – α)100% PARA UNA MEDIA POBLACIONAL μ

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

donde $z_{\alpha/2}$ es el valor z correspondiente a un área $\alpha/2$ en la cola superior de una distribución z normal estándar y

n = tamaño muestral
σ = desviación estándar de la población muestreada

Si σ es desconocida, puede ser aproximada por la desviación estándar muestral s cuando el tamaño muestral sea grande ($n \geq 30$) y el intervalo aproximado de confianza es

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}}$$

Otra forma de hallar el intervalo de confianza de muestra grande para una media poblacional μ es empezar con la estadística

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

que tiene una distribución normal estándar. Si escribimos $z_{\alpha/2}$ como el valor de z con área $\alpha/2$ a su derecha, entonces se puede escribir

$$P\left(-z_{\alpha/2} < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1 - \alpha$$

Esta desigualdad puede reescribirse como

$$-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$-\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < -\mu < -\bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

de modo que

$$P\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

Los términos $\bar{x} - z_{\alpha/2}(\sigma/\sqrt{n})$ y $\bar{x} + z_{\alpha/2}(\sigma/\sqrt{n})$, que son los límites inferior y superior de confianza, son en realidad cantidades aleatorias que dependen de la media muestral $\bar{x}$. Por tanto, en muestreo repetido, el intervalo aleatorio $\bar{x} \pm z_{\alpha/2}(\sigma/\sqrt{n})$, contendrá la media poblacional μ con probabilidad $(1 - \alpha)$.

EJEMPLO 3.6 Un nutriólogo seleccionó una muestra aleatoria de $n = 50$ adultos hombres. Se encontró que el promedio de ingesta diaria de productos lácteos fue de $\bar{x} = 756$ gramos por día, con una desviación estándar de $s = 35$ gramos por día. Use esta información muestral para construir un intervalo de confianza de 95% para la ingesta diaria media de productos lácteos para hombres.

Solución Como el tamaño muestral de $n = 50$ es grande, la distribución de la media muestral $\bar{x}$ está distribuida normalmente en forma aproximada, con media μ y error estándar estimado por $s/\sqrt{n}$. El intervalo de confianza aproximado de 95% es

$$\bar{x} \pm 1.96\left(\frac{s}{\sqrt{n}}\right)$$

$$756 \pm 1.96\left(\frac{35}{\sqrt{50}}\right)$$

$$756 \pm 9.70$$

Por tanto, el intervalo de confianza de 95% para μ es de 746.30 a 765.70 gramos por día.

TIP MI CONSEJO

Un intervalo de confianza de 95% nos dice que, si construyéramos muchos de estos intervalos (todos los cuales tendrían puntos extremos ligeramente diferentes), 95% de ellos encerrarían la media poblacional

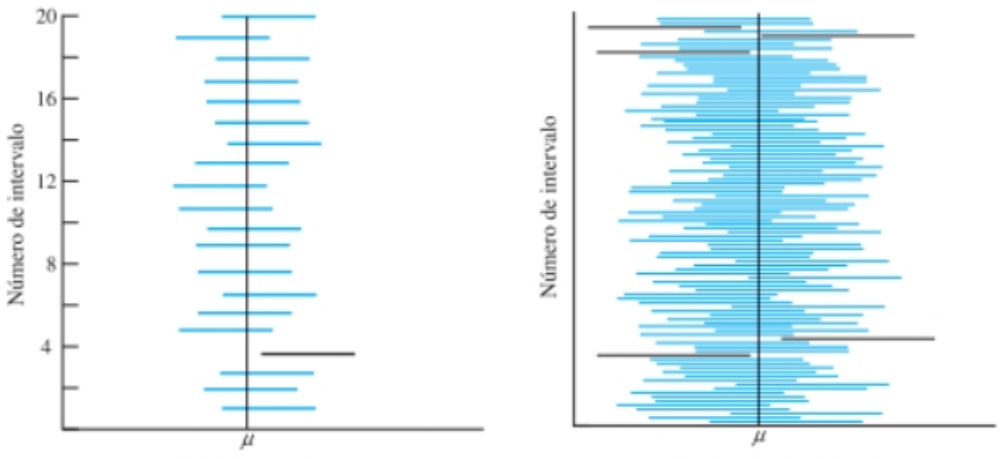
Interpretación del intervalo de confianza

¿Qué significa decir que estamos "95% confiados" que el valor real de la media poblacional μ está dentro de un intervalo determinado? Si construyéramos 20 de esos intervalos, cada uno usando diferente información muestral, nuestros intervalos se verían como los de la figura 3.9a. De los 20 intervalos, podría esperarse que 95% de ellos, o sea 19 de cada 20, funcionaran como se planea y contienen μ dentro de sus límites superior e inferior. Si se construyeran 100 de dichos intervalos (figura 3.9b), se esperaría que alrededor de 95 de ellos se desempeñaran según lo planeado. Recuerde que no se puede estar absolutamente seguro de que algún intervalo particular contenga la media μ. Nunca se sabrá si ese intervalo particular es uno de los 19 que "funcionaron", o si es el intervalo que "faltaba". La confianza en el intervalo estimado proviene del hecho de que cuando se calculan intervalos repetidos, 95% de esos intervalos contendrán μ.

Un buen intervalo de confianza tiene dos características deseables:

- Es tan angosto como es posible. Cuanto más angosto sea el intervalo, más exactamente se habrá localizado el parámetro estimado.
- Tiene un coeficiente de confianza grande, cercano a 1. Cuanto mayor sea el coeficiente de confianza, es más probable que el intervalo contenga el parámetro estimado.

FIGURA 3.9
Interpretación de los intervalos de confianza



(a) 20 intervalos (b) 100 intervalos

MI APPLET EN LÍNEA

Interpretación de intervalos de confianza

EJEMPLO 3.7

Construya un intervalo de confianza de 99% para la ingesta diaria media de productos lácteos para los hombres adultos del ejemplo 3.6.

Solución Para cambiar el nivel de confianza a .99 se debe hallar el valor apropiado de la z normal estándar que pone el área $(1 - \alpha) = .99$ en el centro de la curva. Este valor, con área de cola $\alpha/2 = .005$ a su derecha, se encuentra de la tabla 3.2 como $z = 2.58$ (véase la figura 3.10). El intervalo de confianza de 99% es entonces

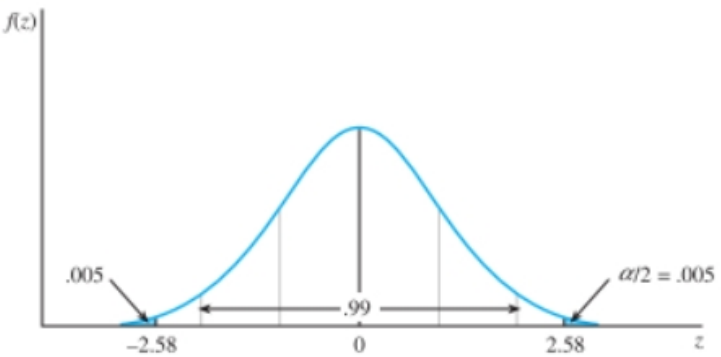
$$\bar{x} \pm 2.58\left(\frac{s}{\sqrt{n}}\right)$$

$$756 \pm 2.58(4.95)$$

$$756 \pm 12.77$$

es decir, 743.23 a 768.77 gramos por día. Este intervalo de confianza es *más ancho* que el intervalo de confianza de 95% del ejemplo 3.6.

FIGURA 3.10
Valores estándar normales para un intervalo de confianza de 99%



TIP MI CONSEJO

Área de la cola derecha	Valor z
.05	1.645
.025	1.96
.01	2.33
.005	2.58

El ancho aumentado es necesario para incrementar la confianza, ¡igual que como se desearía un anillo más ancho en su lazo para asegurarse de lazar el poste de una cerca! La única forma de *aumentar la confianza* sin aumentar el ancho del intervalo es *aumentar el tamaño muestral*, n.

El error estándar de $\bar{x}$,

$$\text{SE} = \frac{\sigma}{\sqrt{n}}$$

MI APPLET EN LÍNEA
Explorar intervalos de confianza

mide la variabilidad o dispersión de los valores de $\bar{x}$. Cuanto más variables sean los datos poblacionales, medidos por σ, más variable será $\bar{x}$ y el error estándar será más grande. Por otra parte, si se aumenta el tamaño muestral n, habrá más información para estimar μ. Las estimaciones deben caer más cerca de μ y el error estándar será más pequeño.

Los intervalos de confianza de los ejemplos 3.6 y 3.7 son aproximados porque se sustituyó s como una aproximación para σ. Esto es, en lugar de que el coeficiente de confianza sea .95, el valor especificado en el ejemplo, el valor verdadero del coeficiente puede ser .92, .94 o .97. Pero esta discrepancia es de poco interés desde un punto de vista práctico; en lo que se refiere a la "confianza" del usuario, hay poca diferencia entre estos coeficientes de confianza. Casi todos los estimadores que se emplean en estadística dan intervalos de confianza aproximados, porque las suposiciones sobre las que están basados no se satisfacen exactamente. Habiendo

visto este punto, no continuaremos refiriéndonos a intervalos de confianza como "aproximados". Es de poco interés práctico mientras el coeficiente real de confianza sea cercano al valor especificado.

Intervalo de confianza de muestra grande para una proporción poblacional *p*

Muchos experimentos de investigación o estudios muestrales tienen como objetivo la estimación de la proporción de personas u objetos de un grupo grande, que posean cierta característica. Veamos algunos ejemplos:

- La proporción de ventas que se puede esperar en un gran número de contactos con clientes
- La proporción de semillas que germinan
- La proporción de votantes "probables" que planean votar por un candidato político particular

Cada uno es un ejemplo práctico del experimento binomial y el parámetro a estimarse es la proporción binomial p.

Cuando el tamaño muestral es grande, la proporción muestral,

$$\hat{p} = \frac{x}{n} = \frac{\text{Número total de éxitos}}{\text{Número total de intentos}}$$

es el mejor estimador puntual para la proporción poblacional p. Como su distribución muestral es aproximadamente normal, con media p y error estándar $\text{SE} = \sqrt{pq/n}$, $\hat{p}$ puede usarse para construir un intervalo de confianza de acuerdo con el método general dado en esta sección.

TIP MI CONSEJO

Área de la cola derecha	Valor z
.05	1.645
.025	1.96
.01	2.33
.005	2.58

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE $(1 - \alpha)100\%$ PARA UNA PROPORCIÓN POBLACIONAL p

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{pq}{n}}$$

donde $z_{\alpha/2}$ es el valor z correspondiente a un área de $\alpha/2$ en la cola derecha de una distribución normal estándar z. Como p y q son incógnitas, se estiman con el uso de los mejores estimadores puntuales: $\hat{p}$ y $\hat{q}$. El tamaño muestral se considera grande cuando la aproximación normal a la distribución binomial es adecuada, es decir, cuando $n\hat{p} > 5$ y $n\hat{q} > 5$.

EJEMPLO 3.8

Una muestra aleatoria de 985 "probables" electores, es decir los que probablemente voten en la próxima elección, fueron encuestados durante un maratón telefónico realizado por el Partido Republicano. De ellos, 592 indicaron que tenían la intención de votar por la candidata republicana en la próxima elección. Construya un intervalo de confianza de 90% para p, la proporción de electores probables de la población que tienen la intención de votar por la candidata republicana. Con base en esta información, ¿se puede concluir que la candidata ganará la elección?

Solución La estimación puntual para p es

$$\hat{p} = \frac{x}{n} = \frac{592}{985} = .601$$

y el error estándar estimado es

$$\sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(.601)(399)}{985}} = .016$$

El valor z para un intervalo de confianza de 90% es el valor que tiene área de $\alpha/2 = .05$ en la cola superior de la distribución z, o $z_{.05} = 1.645$ de la tabla 3.2. El intervalo de confianza del 90% para p es entonces

$$\hat{p} \pm 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$.601 \pm .026$$

Es decir, $.575 < p < .627$. Se estima que el porcentaje de probables electores que tienen intención de votar por la candidata republicana es entre 57.5% y 62.7%. ¿La candidata ganará la elección? Suponiendo que ella necesita más del 50% de los votos para ganar, y como los límites superior e inferior de confianza exceden de este valor mínimo, se puede decir con 90% de confianza que la candidata ganará.

Hay algunos problemas, no obstante, con este tipo de encuesta muestral. ¿Qué pasa si los electores que se consideran a sí mismos "probables para votar" en realidad no van a las casillas? ¿Qué pasa si un elector cambia de idea entre ahora y el día de la elección? ¿Qué pasa si un elector entrevistado no responde honestamente cuando el trabajador de la campaña le hace preguntas? El intervalo de confianza de 90% que ha construido le da 90% de confianza sólo si ha seleccionado una *muestra aleatoria de la población de interés.* Ya no se puede estar seguro de "90% de confianza" si su muestra es sesgada, o ¡si la población de respuestas de votantes cambia antes del día de la elección!

Es posible que usted haya observado que el estimador puntual con su 95% de margen de error se ve muy semejante a un intervalo de confianza de 95% para el mismo parámetro. Esta cercana relación existe para casi todos los parámetros estimados en este libro, pero no es verdadera en general. A veces el mejor estimador puntual para un parámetro *no* cae en la mitad del mejor intervalo de confianza; el mejor intervalo de confianza puede no ser siquiera una función del mejor estimador puntual. Aun cuando ésta es una distinción teórica, debe recordarse que hay una diferencia entre estimación puntual y estimación de intervalo, y que la elección entre las dos depende de la preferencia del experimentador.

3.3 ESTIMACIÓN DE LA DIFERENCIA DE DOS MEDIAS

Un problema de igual importancia que la estimación de una sola media poblacional μ, para una población cuantitativa, es la comparación de dos medias poblacionales. Usted puede hacer comparaciones como éstas:

- Las calificaciones promedio del examen de admisión para la facultad de medicina (MCAT) para estudiantes cuya especialización fuera bioquímica, y para aquellos cuya especialización fuera biología
- Las producciones promedio en una planta química que usa materias primas suministradas por dos proveedores diferentes
- El promedio de diámetros de tallos de plantas cultivadas con dos tipos diferentes de nutrientes

Para cada uno de estos ejemplos, hay dos poblaciones: la primera con media y varianza μ_1 y σ_1^2, y la segunda con media y varianza μ_2 y σ_2^2. Una muestra aleatoria de n_1 mediciones se saca de la población 1 y una segunda muestra aleatoria de tamaño n_2 se saca de manera independiente de la población 2. Por último, las estimaciones de los parámetros poblacionales se calculan a partir de los datos muestrales usando los estimadores $\bar{x}_1$, s_1^2, $\bar{x}_2$ y s_2^2, como se presenta en la tabla 3.3.

TABLA 3.3 **Muestras de dos poblaciones cuantitativas**

	Población 1	Población 2
Media	μ_1	μ_2
Varianza	σ_1^2	σ_2^2

	Muestra 1	Muestra 2
Media	$\bar{x}_1$	$\bar{x}_2$
Varianza	s_1^2	s_2^2
Tamaño de la nuestra	n_1	n_2

Intuitivamente, la diferencia entre dos medias muestrales daría la máxima información acerca de la diferencia real entre dos medias poblacionales y éste es de hecho el caso. El mejor estimador puntual de la diferencia $(\mu_1 - \mu_2)$ entre las medias poblacionales es $(\bar{x}_1 - \bar{x}_2)$. La distribución muestral de este estimador no es difícil de deducir, pero la expresamos aquí sin demostración.

PROPIEDADES DE LA DISTRIBUCIÓN MUESTRAL DE $(\bar{x}_1 - \bar{x}_2)$, LA DIFERENCIA ENTRE DOS MEDIAS MUESTRALES

Cuando muestras aleatorias independientes de n_1 y n_2 observaciones han sido seleccionadas de poblaciones con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2, respectivamente, la distribución muestral de la diferencia $(\bar{x}_1 - \bar{x}_2)$ tiene las siguientes propiedades:

1. La diferencia de medias muestrales es un estimador puntual de la diferencia de medias poblacionales, y el error estándar es:

   $$\text{SE} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

   que se puede estimar como

   $$\text{SE} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$ cuando los tamaños muestrales son grandes.

2. **Si las poblaciones muestreadas están distribuidas normalmente**, entonces la distribución muestral de $(\bar{x}_1 - \bar{x}_2)$ tiene una distribución normal **exacta**, cualquiera que sea el tamaño muestral.
3. **Si las poblaciones muestreadas no están distribuidas normalmente**, entonces la distribución muestral de $(\bar{x}_1 - \bar{x}_2)$ tiene una distribución **aproximadamente** normal cuando n_1 y n_2 son ambas de 30 o más, debido al teorema del límite central.

Como $(\mu_1 - \mu_2)$ es la media de la distribución muestral, se deduce que $(\bar{x}_1 - \bar{x}_2)$ es un estimador insesgado de $(\mu_1 - \mu_2)$ con una distribución aproximadamente normal cuando n_1 y n_2 son grandes. Esto es, el estadístico

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

tiene una distribución z normal aproximadamente estándar y los procedimientos generales de la sección anterior se pueden usar para construir estimaciones puntuales y de intervalo. Aun cuando la elección entre estimación puntual y de intervalo depende de la preferencia personal del usuario, casi todos los experimentadores eligen construir intervalos de confianza para problemas de dos muestras. Las fórmulas apropiadas para ambos métodos se dan a continuación.

ESTIMACIÓN PUNTUAL DE $(\mu_1 - \mu_2)$ DE MUESTRA GRANDE

Estimador puntual: $(\bar{x}_1 - \bar{x}_2)$

95% margen de error: ± 1.96 SE $= \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

TIP MI CONSEJO

Área de la cola derecha	Valor z
.05	1.645
.025	1.96
.01	2.33
.005	2.58

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE DE $(1 - \alpha)100\%$ PARA $(\mu_1 - \mu_2)$

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

EJEMPLO 3.9

Las resistencias al desgaste de dos tipos de llantas para automóvil se compararon en muestras de pruebas en camino de $n_1 = n_2 = 100$ llantas para cada tipo y se registró el número de millas recorridas hasta su deterioro, definido como una cantidad específica de desgaste de la llanta. Los resultados de la prueba se muestran en la tabla 3.4. Estime $(\mu_1 - \mu_2)$, la diferencia en la media de millas hasta el completo desgaste, usando un intervalo de confianza de 99%. ¿Hay diferencia en el promedio de calidad de desgaste para los dos tipos de llantas?

TABLA 3.4 Resumen de datos muestrales para dos tipos de llantas

Llanta 1	Llanta 2
$\bar{x}_1 = 26\,400$ millas	$\bar{x}_2 = 25\,100$ millas
$s_1^2 = 1\,440\,000$	$s_2^2 = 1\,960\,000$

Solución La estimación puntual de $(\mu_1 - \mu_2)$ es

$$(\bar{x}_1 - \bar{x}_2) = 26{,}400 - 25{,}100 = 1300 \text{ millas}$$

y el error estándar de $(\bar{x}_1 - \bar{x}_2)$ se estima como

$$\text{SE} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{1\,440\,000}{100} + \frac{1\,960\,000}{100}} = 184.4 \text{ millas}$$

El intervalo de confianza de 99% se calcula como

$$(\bar{x}_1 - \bar{x}_2) \pm 2.58\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$1300 \pm 2.58(184.4)$$

$$1300 \pm 475.8$$

es decir, $824.2 < (\mu_1 - \mu_2) < 1775.8$. La diferencia en el promedio de millas hasta el completo desgaste para los dos tipos de llantas se estima que está entre el límite inferior de confianza LCL = 824.2 y el límite superior de confianza UCL = 1 775.8 millas.

Con base en este intervalo de confianza, ¿se puede concluir que hay una diferencia en el promedio de millas hasta el completo desgaste para los dos tipos de llantas? Si no hubiera diferencia en las dos medias poblacionales, entonces μ_1 y μ_2 serían iguales y $(\mu_1 - \mu_2) = 0$. Si observamos el intervalo de confianza construido, se verá que 0 no es uno de los posibles valores para $(\mu_1 - \mu_2)$. Por tanto, no es probable que las medias sean iguales; se puede concluir que hay una diferencia en el promedio de millas hasta el completo desgaste para los dos tipos de llantas. El intervalo de confianza ha permitido *tomar una decisión* acerca de la igualdad de las dos medias poblacionales.

TIP MI CONSEJO

Si 0 no es el intervalo, *se puede* concluir que hay una diferencia en las medias poblacionales

EJEMPLO 3.10

El científico del ejemplo 3.6 se preguntaba si había diferencia en el promedio de ingesta diaria de productos lácteos entre hombres y mujeres. Tomó una muestra de $n_1 = 50$ hombres adultos y $n_2 = 50$ mujeres adultas y registró sus ingestas diarias de productos lácteos en gramos por día. En la tabla 3.5 se presenta un resumen de sus resultados muestrales. Construya un intervalo de confianza de 95% para la diferencia en el promedio de ingestas diarias de productos lácteos para hombres y mujeres. ¿Se puede concluir que hay una diferencia en el promedio de ingestas diarias para hombres y mujeres?

TABLA 3.5 Valores muestrales para ingestas diarias de productos lácteos

	Hombres	Mujeres
Tamaño muestral	50	50
Media muestral	756	762
Desviación estándar muestral	35	30

Solución El intervalo de confianza se construye usando un valor de z con área de cola $\alpha/2 = .025$ a su derecha, esto es, $z_{.025} = 1.96$. Usando las desviaciones muestrales estándar para aproximar las desviaciones estándar poblacionales desconocidas, el intervalo de 95% de confianza es

$$(\bar{x}_1 - \bar{x}_2) \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(756 - 762) \pm 1.96\sqrt{\frac{35^2}{50} + \frac{30^2}{50}}$$

$$-6 \pm 12.78$$

o bien, $-18.78 < (\mu_1 - \mu_2) < 6.78$. Veamos los posibles valores para $(\mu_1 - \mu_2)$ del intervalo de confianza. Es posible que la diferencia $(\mu_1 - \mu_2)$ pudiera ser negativa (lo cual indica que el

promedio para mujeres excede del promedio para hombres), ser positiva (lo cual indica que los hombres tienen el promedio más alto) o ser 0 (lo cual indica que no hay diferencia entre los promedios). Con base en esta información, *no estaríamos dispuestos a concluir* que hay una diferencia en el promedio de ingestas diarias de productos lácteos para hombres y mujeres.

Los ejemplos 3.9 y 3.10 merecen más comentarios respecto a usar estimaciones muestrales en lugar de parámetros desconocidos. La distribución muestral de

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

tiene una distribución normal estándar para todos los tamaños muestrales cuando ambas poblaciones muestreadas son normales, y una distribución normal estándar *aproximada* cuando las poblaciones muestreadas no sean normales pero los tamaños muestrales sean grandes (≥ 30). Cuando σ_1^2 y σ_2^2 no se conocen y son estimadas por las estimaciones muestrales s_1^2 y s_2^2, la estadística resultante todavía tendrá una distribución normal estándar aproximada cuando los tamaños muestrales sean grandes. El comportamiento de esta estadística cuando las varianzas poblacionales son desconocidas, y los tamaños muestrales sean pequeños, se estudiará en el capítulo 4.

3.4 ESTIMACIÓN DEL PARÁMETRO DE UNA POBLACIÓN BINOMIAL, ESTIMACIÓN DE LA DIFERENCIA ENTRE DOS PROPORCIONES

Una simple extensión de la estimación de una proporción binomial p es la estimación de la diferencia entre dos proporciones binomiales. Se pueden hacer comparaciones como éstas:

- La proporción de artículos defectuosos manufacturados en dos líneas de producción
- La proporción de votantes mujeres y hombres que están a favor de una enmienda por la igualdad de derechos
- Los porcentajes de germinación de semillas no tratadas y semillas tratadas con un fungicida

Estas comparaciones pueden hacerse con la diferencia $(p_1 - p_2)$ entre dos proporciones binomiales, p_1 y p_2. Muestras aleatorias independientes formadas por n_1 y n_2 intentos se sacan de poblaciones 1 y 2, respectivamente, y se calculan las estimaciones muestrales $\hat{p}_1$ y $\hat{p}_2$. El estimador insesgado de la diferencia $(p_1 - p_2)$ es la diferencia muestral $(\hat{p}_1 - \hat{p}_2)$.

PROPIEDADES DE LA DISTRIBUCIÓN MUESTRAL DE LA DIFERENCIA $(\hat{p}_1 - \hat{p}_2)$ ENTRE DOS PROPORCIONES MUESTRALES

Suponga que las muestras aleatorias independientes de las observaciones n_1 y n_2 han sido seleccionadas de poblaciones binomiales con parámetros p_1 y p_2, respectivamente. La distribución muestral de la diferencia entre proporciones muestrales

$$(\hat{p}_1 - \hat{p}_2) = \left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right)$$

tiene estas propiedades:

1. La media de $(p_1 - p_2)$ es

   $$p_1 - p_2$$

   y el error estándar es

   $$\text{SE} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

   que se estima como

   $$\text{SE} = \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

2. La distribución muestral de $(p_1 - p_2)$ puede ser aproximada por una distribución normal cuando n_1 y n_2 son grandes, debido al teorema del límite central.

Aun cuando el margen de una proporción individual es de 0 a 1, la diferencia entre dos proporciones va de -1 a 1. Para usar una distribución normal para aproximar la distribución de $(\hat{p}_1 - \hat{p}_2)$, tanto $\hat{p}_1$ como $\hat{p}_2$ deben ser aproximadamente normales; esto es, $n_1\hat{p}_1 > 5$, $n_1\hat{q}_1 > 5$ y $n_2\hat{p}_2 > 5$, $n_2\hat{q}_2 > 5$.

Las fórmulas apropiadas para estimación puntual y de intervalo se dan a continuación.

ESTIMACIÓN PUNTUAL DE MUESTRA GRANDE DE $(p_1 - p_2)$

Estimador puntual: $(\hat{p}_1 - \hat{p}_2)$

95% de margen de error: $\pm 1.96\ \text{SE} = \pm 1.96\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$

UN INTERVALO DE CONFIANZA DE MUESTRA GRANDE $(1 - \alpha)100\%$ PARA $(p_1 - p_2)$

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Suposición: n_1 y n_2 deben ser suficientemente grandes para que la distribución muestral de $(\hat{p}_1 - \hat{p}_2)$ pueda ser aproximada por una distribución normal, es decir, si $n_1\hat{p}_1$, $n_1\hat{q}_1$, $n_2\hat{p}_2$ y $n_2\hat{q}_2$ son todas mayores que 5.

EJEMPLO 3.11

La propuesta de un bono para la construcción de una escuela será enviada a los votantes en la siguiente elección municipal. Una parte importante del dinero derivado de esta emisión de bonos se empleará en construir escuelas en una zona de rápido desarrollo de la ciudad y lo demás se usará para renovar y actualizar los edificios escolares del resto de ésta. Para evaluar la viabilidad de la propuesta de un bono, a una muestra aleatoria de $n_1 = 50$ residentes de la zona de rápido desarrollo y $n_2 = 100$ de las otras partes de la ciudad, se les preguntó si piensan votar por la propuesta. Los resultados se tabulan en la tabla 3.6.

TABLA 3.6 **Valores muestrales para opinión sobre propuesta de bono**

	Sección de desarrollo	Resto de la ciudad
Tamaño muestral	50	100
Número a favor de propuesta	38	65
Proporción a favor de propuesta	.76	.65

1. Estime la diferencia en las proporciones verdaderas a favor de la propuesta de bono con un 99% de intervalo de confianza.
2. Si ambas muestras se agrupan en una muestra de tamaño $n = 150$, con 103 a favor de la propuesta, dé una estimación puntual de la proporción de residentes de la ciudad que votarán por la propuesta del bono. ¿Cuál es el margen de error?

Solución

1. La mejor estimación puntual de la diferencia $(p_1 - p_2)$ está dada por

   $$(\hat{p}_1 - \hat{p}_2) = .76 - .65 = .11$$

   y el error estándar de $(\hat{p}_1 - \hat{p}_2)$ se estima como

   $$\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} = \sqrt{\frac{(.76)(.24)}{50} + \frac{(.65)(.35)}{100}} = .0770$$

   Para un intervalo de confianza de 99%, $z_{.005} = 2.58$ y el intervalo aproximado de confianza de 99% se encuentra como

   $$(\hat{p}_1 - \hat{p}_2) \pm z_{.005}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

   $$.11 \pm (2.58)(.0770)$$

   $$.11 \pm .199$$

   o bien, (–.089, .309). Como este intervalo contiene el valor $(p_1 - p_2) = 0$, es posible que $p_1 = p_2$, lo cual implica que puede no haber diferencia en las proporciones a favor del asunto del bono en las dos secciones de la ciudad.

2. Si no hay diferencia en las dos proporciones, entonces las dos muestras no son realmente diferentes y podrían combinarse bien para obtener una estimación total de la proporción de los residentes de la ciudad que votarán por el asunto del bono. Si ambas muestras se agrupan, entonces $n = 150$ y

   $$\hat{p} = \frac{103}{150} = .69$$

   Por tanto, la estimación puntual del valor total de p es .69, con un margen de error dado por

   $$\pm 1.96\sqrt{\frac{(.69)(.31)}{150}} = \pm 1.96(.0378) = \pm .074$$

   Observe que $.69 \pm .074$ produce el intervalo .62 a .76, que incluye sólo proporciones mayores que .5. Por tanto, si las actitudes de los votantes no cambian de manera adversa antes de la elección, la propuesta del bono debe aprobarse por una mayoría razonable.

3.5 SELECCIÓN DEL TAMAÑO DE LA MUESTRA

Diseñar un experimento es en esencia un plan para comprar cierta cantidad de información. Así como el precio que se paga por un juego de video varía dependiendo de dónde y cuándo se compra, el precio de información estadística varía dependiendo de cómo y dónde se recolecta la información. Al igual que cuando se compra cualquier producto, se debe comprar tanta información estadística como sea posible por el mínimo costo posible.

La cantidad total de información relevante en una muestra es controlada por dos factores:

- El **plan muestral** o **diseño experimental**: el procedimiento para recolectar la información
- El **tamaño muestral** n: la cantidad de información recolectada

Se puede incrementar la cantidad de información recolectada al *aumentar* el tamaño muestral o quizás al *cambiar* el tipo de plan muestral o diseño experimental que se utilice. Trataremos el plan muestral más sencillo, es decir, el muestreo aleatorio de una población relativamente grande, y nos concentraremos en las formas para elegir el tamaño muestral n necesario para comprar una cantidad determinada de información.

Un investigador hace poco progreso al planear un experimento antes de encontrar el problema del tamaño muestral. **¿Cuántas mediciones deben incluirse en la muestra?** ¿Cuánta información desea comprar el investigador? Con el fin de responder estas preguntas, el investigador debe especificar primero:

- La confiabilidad que desea lograr, y
- La precisión necesaria para su estimación

En un problema de estimación estadística, la precisión de la estimación se mide por el *margen de error* o el *ancho del intervalo de confianza*, los cuales tienen una confiabilidad especificada. En vista que ambas medidas son una función del tamaño de la muestra, especificar la confiabilidad y la precisión permite determinar el tamaño necesario de la muestra.

Por ejemplo, suponga que se desea estimar el promedio de producción diaria μ de un proceso químico y se necesita que el margen de error sea menor que 4 toneladas. Esto significa que, aproximadamente 95% de las veces en muestreo repetido, la "confiabilidad", la distancia entre la media muestral $\bar{x}$ y la media poblacional μ será menor que 1.96 SE. Se desea que esta cantidad sea menor que 4 (la "precisión"). Esto es,

$$1.96\,\text{SE} < 4 \quad \text{o} \quad 1.96\left(\frac{\sigma}{\sqrt{n}}\right) < 4$$

Despejando n, resulta

$$n > \left(\frac{1.96}{4}\right)^2 \sigma^2 \quad \text{o} \quad n > .24\sigma^2$$

Si se conoce σ, la desviación estándar poblacional, puede sustituirse su valor en la fórmula y despejar n. Si σ es desconocida, que suele ser el caso, puede usarse la mejor aproximación disponible:

- Una estimación s se obtiene de una muestra previa
- Una estimación del rango basada en el conocimiento de las mediciones máximas y mínimas posibles: $\sigma \approx \text{Rango}/4$

Para este ejemplo, suponga que un estudio previo del proceso químico produjo una desviación muestral estándar de $s = 21$ toneladas. Entonces

$$n > .24\sigma^2 = .24(21)^2 = 105.8$$

Usando una muestra de tamaño $n = 106$ o mayor, se puede estar razonablemente seguro (con probabilidad aproximadamente igual a .95) que su estimación del promedio de producción estará a no más de ± 4 toneladas del promedio real de producción.

La solución $n = 106$ es sólo aproximada porque tenía que usarse un valor aproximado de σ para calcular el error estándar de la media. Aun cuando esto puede ser molesto, es el mejor método existente para seleccionar el tamaño muestral ¡y es ciertamente mejor que adivinar!

A veces los investigadores requieren una confiabilidad diferente, o nivel de confianza distinto a 95% de confianza especificado por el margen de error. En este caso, la mitad del ancho del intervalo de confianza da la medida de precisión para su estimación; esto es, el límite B en el error de su estimación es

$$z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) < B$$

Este método de seleccionar el tamaño muestral se puede usar para los cuatro procedimientos de estimación presentados en este capítulo. El procedimiento general se describe a continuación.

NECESITO SABER...

Cómo seleccionar el tamaño muestral

Determine el parámetro a estimarse y el error estándar de su estimador puntual. A continuación proceda como sigue:

1. Elija B, el límite en el error de su estimación y un coeficiente de confianza $(1 - \alpha)$.
2. Para un problema de una muestra, de esta ecuación despeje el tamaño muestral n:

   $$z_{\alpha/2} \times (\text{error estándar del estimador}) \leq B$$

   donde $z_{\alpha/2}$ es el valor de z que tiene un área $\alpha/2$ a su derecha.
3. Para un problema de dos muestras, haga $n_1 = n_2 = n$ y resuelva la ecuación del paso 2.

[NOTA: Para casi todos los estimadores, el error estándar es una función del tamaño muestral n.]

EJEMPLO 3.12

Los productores de tubo de plástico de polivinilo desean tener un suministro de tubos suficiente para satisfacer las necesidades del mercado. Desean encuestar mayoristas que compren tubos de polivinilo para estimar la proporción que planea aumentar sus compras el año próximo. ¿Qué tamaño muestral se requiere si desean que su estimación se encuentre a no más de .04 de la proporción real con probabilidad igual a .90?

Solución Para este ejemplo particular, el límite B en el error de la estimación es .04. Como el coeficiente de confianza es $(1 - \alpha) = .90$, α debe ser igual a .10 y $\alpha/2$ es .05. El valor z correspondiente a un área igual a .05 en la cola superior de la distribución z es $z_{.05} = 1.645$. Entonces se requiere

$$1.645 \text{ SE} = 1.645\sqrt{\frac{pq}{n}} \leq .04$$

Para despejar n de esta ecuación, debe sustituirse un valor aproximado de p en la ecuación. Si se desea estar seguro que la muestra es suficientemente grande, debe usarse $p = .5$ (sustituir $p = .5$ dará la máxima solución posible para n porque el máximo valor de pq se presenta cuando $p = q = .5$). Entonces

$$1.645\sqrt{\frac{(.5)(.5)}{n}} \leq .04$$

o bien

$$\sqrt{n} \geq \frac{(1.645)(.5)}{.04} = 20.56$$

$$n \geq (20.56)^2 = 422.7$$

Por tanto, los productores deben incluir al menos 423 mayoristas en su encuesta si desean estimar la proporción p correcta a no más de .04.

EJEMPLO 3.13

Un director de personal desea comparar la eficacia de dos métodos de capacitar empleados industriales a realizar cierta operación de ensamble. Un número de empleados se divide en dos grupos iguales: el primero recibiendo capacitación del método 1, y el segundo, el método de capacitación 2. Cada uno efectuará la operación de ensamble y se registrará el tiempo de ensamblado. Se espera que las mediciones para ambos grupos tendrán una variación de alrededor de 8 minutos. Para que la estimación de la diferencia en tiempos medios de ensamble sea correcta a no más de 1 minuto de variación, con una probabilidad igual a .95, ¿cuántos trabajadores deben estar incluidos en cada grupo de capacitación?

Solución En vista que se está estimando la diferencia entre dos medias, el error estándar de la estimación es $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$. El límite es $B = 1$ minuto, de modo que se requiere

$$1.96\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq 1$$

Como se desea usar dos grupos iguales, se puede hacer $n_1 = n_2 = n$ y obtener la ecuación

$$1.96\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}} \leq 1$$

Como ya observamos, la variabilidad (intervalo) de cada método de ensamble es aproximadamente igual y, por lo tanto, $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Como el intervalo, igual a 8 minutos, es aproximadamente igual a 4σ, tenemos

$$4\sigma \approx 8 \quad \text{o bien} \quad \sigma \approx 2$$

Sustituyendo este valor por σ_1 y σ_2 en la ecuación anterior, se obtiene

$$1.96\sqrt{\frac{(2)^2}{n} + \frac{(2)^2}{n}} \leq 1$$

$$1.96\sqrt{\frac{8}{n}} \leq 1$$

$$\sqrt{n} \geq 1.96\sqrt{8}$$

Resolviendo, tenemos $n \geq 31$. Por tanto, cada grupo debe contener al menos $n = 31$ trabajadores.

La tabla 3.7 contiene un resumen de las fórmulas empleadas para determinar los tamaños muestrales requeridos para estimación, con un límite dado en el error de la estimación o ancho de intervalo de confianza W ($W = 2B$). Observe que para estimar p, la fórmula del tamaño muestral usa $\sigma^2 = pq$, en tanto que para estimar $(p_1 - p_2)$, la fórmula del tamaño muestral usa $\sigma_1^2 = p_1q_1$ y $\sigma_2^2 = p_2q_2$.

TABLA 3.7 **Fórmulas para tamaño muestral**

Parámetro	Estimador	Tamaño muestral	Suposiciones
μ	$\bar{x}$	$n \geq \frac{z_{\alpha/2}^2 \sigma^2}{B^2}$	
$\mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2$	$n \geq \frac{z_{\alpha/2}^2 (\sigma_1^2 + \sigma_2^2)}{B^2}$	$n_1 = n_2 = n$
p	$\hat{p}$	$n \geq \frac{z_{\alpha/2}^2 pq}{B^2}$ 0 $n \geq \frac{(.25) z_{\alpha/2}^2}{B^2}$	 $p = .5$
$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	$n \geq \frac{z_{\alpha/2}^2 (p_1q_1 + p_2q_2)}{B^2}$ 0 $n \geq \frac{2(.25) z_{\alpha/2}^2}{B^2}$	$n_1 = n_2 = n$ $n_1 = n_2 = n$ y $p_1 = p_2 = .5$

3.6 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

1. Exprese el teorema del límite central. ¿De qué valor es el teorema del límite central en estimación estadística de muestra grande?

2. Una muestra aleatoria de $n = 64$ observaciones tiene una media de $\bar{x}$ y una desviación estándar $s = 3.9$.

a. Dé la estimación puntual de la media poblacional μ y encuentre el margen de error para su estimación

b. Encuentre un intervalo de confianza de 90% para μ. ¿Qué significa "90% confiado"?

c. Encuentre un límite inferior de confianza de 90% para la media poblacional μ. ¿Por qué este límite es diferente del límite inferior de confianza de la parte **b**?

d. ¿Cuántas observaciones son necesarias para estimar μ a no más de .5, con probabilidad igual a .95?

3. Se seleccionaron muestras aleatorias independientes de $n_1 = 40$ y $n_2 = 80$ observaciones de las poblaciones binomiales 1 y 2, respectivamente. El número de éxitos en las dos muestras fueron $x_1 = 17$ y $x_2 = 23$. Encuentre un intervalo de confianza de 99% para la diferencia entre las dos proporciones poblacionales binomiales. Interprete este intervalo.

4. **Cocina étnica** Grupos étnicos en Estados Unidos compran diferentes cantidades de diversos productos alimenticios debido a su cocina étnica. A una investigadora interesada en la segmentación de mercado para familias asiáticas e hispanoestadounidenses le gustaría estimar la proporción de familias que seleccionan ciertas marcas de varios productos. Si la investigadora desea que estas

estimaciones se encuentren a no más de .03 con probabilidad de .95, ¿cuántas familias debe incluir en las muestras? Suponga que los tamaños muestrales son iguales.

5. Fumar y presión sanguínea Se realizó un experimento para estimar el efecto de fumar sobre la presión sanguínea de un grupo de 35 fumadores. Se obtuvo la diferencia para cada participante al tomar la diferencia en las lecturas de presión sanguínea al principio del experimento y otra vez cinco años más adelante. El aumento de la media muestral, medido en milímetros de mercurio, fue de $\bar{x} = 9.7$. La desviación estándar muestral fue $s = 5.8$. Estime el aumento medio en presión sanguínea que se esperaría para fumadores de cigarrillos en el espacio de tiempo indicado por el experimento. Encuentre el margen de error. Describa la población asociada con la media que haya estimado.

6. Estaturas Si se supone que las estaturas de hombres están normalmente distribuidas con una desviación estándar de 2.5 pulgadas, ¿qué tan grande debe tomarse una muestra para estar razonablemente seguro (probabilidad .95) de que la media muestral no difiere de la media verdadera (media poblacional) en más de .50 en valor absoluto?

7. ¡Comida rápida! Aun cuando sabemos que puede no ser buena para la salud, ¡muchos estadounidenses en verdad disfrutan su comida rápida! Una encuesta llevada a cabo por *Pew Research Center*[1] ilustró en forma gráfica la inclinación por comer fuera y, en particular, por ingerir comida rápida:

Fuente: Pew Research Center, pewsocialtrends.org

a. Esta encuesta se basó en "entrevistas telefónicas que se llevaron a cabo con una muestra nacionalmente representativa de 2 250 adultos, con edades de 18 años en adelante, que viven en hogares con teléfono en la porción continental de Estados Unidos". ¿Qué problemas podrían surgir con este tipo de muestreo?

b. ¿Qué tan precisos espera que sean los porcentajes dados en la encuesta en la estimación de los porcentajes reales de la población? (SUGERENCIA: Encuentre el margen de error.)

c. Si desea disminuir su margen de error para que sea $\pm 1\%$, ¿qué tan grande debería ser su muestra?

8. Control de calidad Un ingeniero de control de calidad desea estimar la fracción de defectos en un lote grande de cartuchos de tinta para impresora. De experiencias previas, él sabe que la fracción real de defectos debe estar alrededor de .05. ¿Qué tan grande debe tomar una muestra si él desea estimar la verdadera fracción a no más de .01, usando un intervalo de confianza de 95%?

9. Reincidencia Se utilizó una técnica experimental de rehabilitación en prisioneros liberados. Se demostró que 79 de 121 hombres sometidos a la técnica prosiguieron con vidas útiles y sin delincuencia durante un periodo de tres años después de su liberación. Encuentre un intervalo de confianza de 95% para p, la probabilidad de que un prisionero sometido a la técnica de rehabilitación seguirá una vida sin delincuencia durante al menos tres años después de ser liberados.

[1] "Eating More: Enjoying Less," 19 de abril de 2006, *Social and Demographic Trends, Pew Research Center*, http://pewsocialtrends.org/pubs/309/eating-more-enjoying-less, 3 de septiembre de 2010.

10. Investigación de audiología Se pidió a 70 estudiantes hombres de primer año de universidad que participaran en un estudio para establecer el umbral absoluto de audibilidad. Cada individuo tomó asiento en un cuarto a prueba de sonidos y se le presentó un tono de 150 H a un gran número de niveles de estímulo en un orden aleatorio. A la persona se le indicó que presionara un botón si detectaba el sonido; el experimentador registró el nivel más bajo de estímulo al cual se detectó el tono. La media para el grupo fue de 21.6 dB con $s = 2.1$. Estime el umbral medio absoluto para todos los estudiantes universitarios de primer año y calcule el margen de error.

11. La garrapata roja de cítricos Una entomóloga desea estimar el tiempo promedio de desarrollo de la garrapata roja de cítricos, a no más de .5 de día. De los experimentos previos, se sabe que σ es cercana a los cuatro días. ¿Qué tan grande debe ser la muestra que tome la entomóloga para tener 95% de confianza de su estimación?

EJERCICIOS

TÉCNICAS BÁSICAS

3.1 Explique qué significa "margen de error" en estimación puntual.

3.2 ¿Cuáles son dos características del mejor estimador puntual para un parámetro poblacional?

3.3 Calcule el margen de error al estimar una media poblacional μ para estos valores:

a. $n = 30, \sigma^2 = .2$ **b.** $n = 30, \sigma^2 = .9$

c. $n = 30, \sigma^2 = 1.5$

3.4 Una muestra aleatoria de $n = 900$ observaciones de una población binomial produjo $x = 655$ éxitos. Estime la proporción binomial p y calcule el margen de error.

3.5 Una muestra aleatoria de $n = 75$ observaciones de una población cuantitativa produjo $\bar{x} = 29.7$ y $s^2 = 10.8$. Dé la mejor estimación puntual para la media poblacional μ y calcule el margen de error.

3.6 Encuentre e interprete un intervalo de confianza de 95% para una media poblacional μ para estos valores:

a. $n = 36, \bar{x} = 13.1, s^2 = 3.42$

b. $n = 64, \bar{x} = 2.73, s^2 = .1047$

3.7 Encuentre un intervalo de confianza $(1 - \alpha)100\%$ para una media poblacional μ para estos valores:

a. $\alpha = .01, n = 38, \bar{x} = 34, s^2 = 12$

b. $\alpha = .10, n = 65, \bar{x} = 1049, s^2 = 51$

c. $\alpha = .05, n = 89, \bar{x} = 66.3, s^2 = 2.48$

3.8 Encuentre e interprete un intervalo de confianza de 95% para una media poblacional μ para estos valores:

a. $n = 36, \bar{x} = 13.1, s^2 = 3.42$

b. $n = 64, \bar{x} = 2.73, s^2 = .1047$

3.9 Encuentre un intervalo de confianza $(1 - \alpha)100\%$ para una media poblacional μ para estos valores:

a. $\alpha = .01, n = 38, \bar{x} = 34, s^2 = 12$

b. $\alpha = .10, n = 65, \bar{x} = 1049, s^2 = 51$

c. $\alpha = .05, n = 89, \bar{x} = 66.3, s^2 = 2.48$

3.10 Se seleccionaron muestras aleatorias independientes de tamaño $n_1 = n_2 = 100$ de cada una de dos poblaciones. La media y la desviación estándar para las dos muestras fueron $\bar{x}_1 = 125.2$, $\bar{x}_2 = 123.7$, $s_1 = 5.6$ y $s_2 = 6.8$.

a. Construya un intervalo de confianza de 99% para estimar la diferencia en las dos medias poblacionales.

b. ¿El intervalo de confianza en la parte **a** proporciona suficiente evidencia para concluir que hay una diferencia en las dos medias poblacionales? Explique.

3.11 Se seleccionaron muestras aleatorias independientes de $n_1 = 1265$ y $n_2 = 1688$ observaciones de las poblaciones binomiales 1 y 2, y se observaron $x_1 = 849$ y $x_2 = 910$ éxitos.

a. Encuentre un intervalo de confianza de 99% para la diferencia $(p_1 - p_2)$ de las dos proporciones poblacionales. ¿Qué significa "99% de confianza"?

b. Con base en el intervalo de confianza de la parte **a**, ¿puede concluirse que hay una diferencia en las dos proporciones binomiales? Explique.

3.12 Encuentre un límite superior de confianza de una cola al 90% para la media poblacional μ para estos valores:

a. $n = 40, s^2 = 65, \bar{x} = 75$

b. $n = 100, s = 2.3, \bar{x} = 1.6$

3.13 Muestras aleatorias independientes de $n_1 = n_2 = n$ observaciones son seleccionadas de cada una de las poblaciones 1 y 2. Si desea estimar la diferencia entre las dos medias poblacionales correctas a no más de .17 de variación, con probabilidad igual a .90, ¿qué tan grandes deben ser n_1 y n_2? Suponga que se conocen $\sigma_1^2 \approx \sigma_2^2 \approx 27.8$.

APLICACIONES

3.14 Niños multimedia ¿Nuestros hijos pasan el mismo tiempo disfrutando de actividades al aire libre y jugando con la familia y amigos que las generaciones previas? O bien, ¿nuestros hijos pasan cada vez más tiempo frente a un televisor, computadora y otros equipos multimedia? Una muestra aleatoria de 250 niños entre ocho y 18 años de edad mostró que 170 niños tenían un televisor en su recámara y 120 de ellos tenían también un juego de video.

a. Estime la proporción de todos los niños y adolescentes, de ocho a 18 años, que tienen un televisor en sus recámaras y calcule el margen de error de su estimación.

b. Estime la proporción de todos los niños y adolescentes, de ocho a 18 años, que tienen un juego de video en sus recámaras y calcule el margen de error de su estimación.

3.15 Inmigración ilegal En un sondeo reciente que incluyó preguntas sobre inmigración ilegal en Estados Unidos, y las respuestas federales y estatales al problema, 75% de los $n = 1004$ adultos encuestados pensaban que Estados Unidos no se esforzaba lo suficiente para evitar que los inmigrantes ilegales llegaran a ese país.[1]

a. ¿Cuál es una estimación puntual para la proporción de adultos que creen que Estados Unidos no se esfuerza lo suficiente para evitar que inmigrantes ilegales vayan a ese país? Calcule el margen de error.

b. El sondeo refleja un margen de error de $\pm 3.5\%$. ¿Cómo debería calcularse el margen de error mencionado de modo que pueda aplicarse a todas las preguntas en la encuesta? ¿Es correcto el margen de error mencionado?

	Marriott	Westin	Doubletree
Promedio muestral ($)	150	165	125
Desviación estándar muestral	17.2	22.5	12.8

3.16 Ratas hambrientas En un experimento para evaluar la intensidad del instinto del hambre en ratas, 30 animales previamente entrenados fueron privados de alimento durante 24 horas. Al término de ese periodo, cada rata fue puesta en una jaula donde se les dio alimento si el animal presionaba una palanca. Para cada una, se registró el tiempo en el que continuaba presionando la barra (aun cuando no recibiera alimento). Si los datos dieron una media mues-

tral de 19.3 minutos con una desviación estándar de 5.2 minutos, estime el verdadero tiempo medio y calcule el margen de error.

3.17 Lluvia ácida La lluvia ácida, causada por la reacción de ciertos contaminantes del aire con el agua de lluvia, parece ser un problema creciente en Estados Unidos. La lluvia pura que cae en aire limpio registra un valor pH de 5.7 (el pH es una medida de la acidez: 0 es ácido; 14 es alcalino). Suponga que muestras de agua de 40 lluvias se analizan para el contenido del pH y $\bar{x}$ y s son iguales a 3.7 y .5, respectivamente. Encuentre un intervalo de confianza de 99% para el pH medio en agua de lluvia e interprete el intervalo. ¿Qué suposición debe hacerse para que el intervalo de confianza sea válido?

3.18 Ruido y estrés Para comparar el efecto del estrés en la forma de ruido sobre la capacidad de realizar un trabajo sencillo, 70 personas fueron divididas en dos grupos. El primer grupo de 30 personas actuó como control, en tanto que el segundo grupo de 40 fue el grupo experimental. Aun cuando cada persona realizó el trabajo, las personas del grupo experimental realizaron el trabajo cuando se reproducía música de *rock* a alto volumen. Se registró el tiempo para terminar el trabajo para cada individuo y se obtuvo el siguiente resumen:

	Control	Experimental
n	30	40
$\bar{x}$	15 minutos	23 minutos
s	4 minutos	10 minutos

a. Encuentre un intervalo de confianza de 99% para la diferencia en tiempos medios de terminación para estos dos grupos.

b. Con base en el intervalo de confianza de la parte **a**, ¿hay suficiente evidencia para indicar una diferencia en el tiempo promedio de terminación para los dos grupos? Explique.

3.19 M&M's ¿La compañía Mars, Incorporated, usa la misma proporción de dulces rojos en sus variedades sencilla y de cacahuate? Una muestra aleatoria de 56 M&M's sencillos contenía 12 dulces rojos y otra muestra aleatoria de 32 M&M's de cacahuate contenía ocho dulces rojos.

a. Construya un intervalo de confianza de 95% para la diferencia en las proporciones de dulces rojos para las variedades sencilla y de cacahuate.

b. Con base en el intervalo de confianza de la parte **a**, ¿puede concluirse que hay una diferencia en las proporciones de dulces rojos para las variedades sencilla y de cacahuate? Explique.

3.20 ¿Excedrin o Tylenol? En un estudio para comparar los efectos de dos analgésicos se encontró que, de $n_1 = 200$ personas seleccionadas al azar que usaron el primer analgésico, 93% indicó que alivió su dolor. De $n_2 = 450$ personas seleccionadas al azar que usaron el segundo analgésico, 96% indicó que les alivió el dolor.

a. Encuentre un intervalo de confianza de 99% para la diferencia en las proporciones que experimentan alivio por estos dos analgésicos.

b. Con base en el intervalo de confianza de la parte **a**, ¿hay suficiente evidencia para indicar una diferencia en las proporciones que experimentan alivio para los dos analgésicos? Explique.

3.21 Gastos de operación Una muestra aleatoria de los costos mensuales de operación de una compañía para $n = 36$ meses produjo una media muestral de 5474 dólares y una desviación estándar de 764 dólares. Encuentre un límite superior de confianza de 90% para los gastos mensuales medios de la compañía.

3.22 Inmigración ilegal En el ejercicio 3.15 se presentó una encuesta de investigación llevada a cabo para *ABC News* y el *Washington Post* que incluyó preguntas sobre inmigración ilegal en Estados Unidos, y las respuestas federales y estatales al problema.[1] Suponga que usted diseña un sondeo de este tipo.

a. Explique cómo seleccionaría su muestra. ¿Qué problemas podría encontrar en este proceso? b. Si usted desea estimar el porcentaje de la población que está de acuerdo con una frase particular en su cuestionario de encuesta, correcto a no más de 1% con probabilidad .95, ¿aproximadamente cuántas personas tendrían que ser encuestadas?

Fuente

1. "Immigration," *ABC News/Washington Post Poll*, 3–6 de junio de 2010, http://www.pollingreport.com/immigration.htm, 2 de septiembre de 2010.

4

Inferencia estadística con muestras pequeñas

ZouZou/Shutterstock.com

OBJETIVOS GENERALES

En el capítulo 3 se presentaron los conceptos básicos de estimación estadística de muestra grande, y en el capítulo 5 se estudia la prueba de hipótesis para situaciones prácticas que involucran medias y proporciones poblacionales. Como todas estas técnicas se apoyan en el teorema del límite central para justificar la normalidad de los estimadores y estadísticos de prueba, sólo se aplican cuando las muestras son grandes. Este capítulo complementa las técnicas de muestra grande al presentar las pruebas de muestra pequeña y los intervalos de confianza para medias y varianzas poblacionales. A diferencia de sus similares de muestras grandes, estas técnicas de muestra pequeña requieren que las poblaciones muestreadas sean normales o aproximadamente normales.

ÍNDICE DEL CAPÍTULO

- Comparación de dos varianzas poblacionales
- Distribución *t* de Student
- Inferencias de muestra pequeña respecto a la diferencia en dos medias: muestras aleatorias independientes
- Inferencias de muestra pequeña respecto a una media poblacional
- Inferencias respecto a varianza poblacional
- Prueba de diferencia pareada: muestras dependientes

NECESITO SABER...

Cómo decidir cuál prueba usar

Suponga que usted necesita realizar un experimento para estimar una media poblacional o la diferencia entre dos medias. El proceso de recolectar los datos puede ser muy costoso o lento. Si no es posible recolectar una *muestra grande*, los procedimientos de estimación del capítulo 3 no sirven.

Este capítulo presenta algunos procedimientos estadísticos equivalentes que pueden usarse cuando *el tamaño muestral es pequeño*. Los procedimientos de estimación y prueba comprenden estos parámetros ya conocidos:

- Una sola media poblacional, μ
- La diferencia entre dos medias poblacionales, $(\mu_1 - \mu_2)$
- Una sola varianza poblacional, σ^2
- La comparación de dos varianzas poblacionales, σ_1^2 y σ_2^2

Las pruebas e intervalos de confianza de muestra pequeña para proporciones binomiales se omitirán en nuestro análisis.[†]

4.1 LA DISTRIBUCIÓN *t* DE STUDENT

Al efectuar un experimento destinado a evaluar un proceso nuevo pero muy costoso para producir diamantes sintéticos, usted solamente puede estudiar seis diamantes generados por el proceso. ¿Cómo usar estas seis mediciones para hacer inferencias acerca del promedio de peso μ de los diamantes a partir de este proceso?

Al estudiar la distribución muestral de $\bar{x}$ en el capítulo 2, planteamos estos puntos:

- Cuando la población original muestreada es normal, $\bar{x}$ y $z = (\bar{x} - \mu)/(\sigma/\sqrt{n})$ ambas tienen distribuciones normales, *para cualquier tamaño muestral.*
- Cuando la población original muestreada *no* es normal, $\bar{x}$, $z = (\bar{x} - \mu)/(\sigma/\sqrt{n})$, y $z \approx (\bar{x} - \mu)/(s/\sqrt{n})$ todas tienen distribuciones aproximadamente normales, si el tamaño muestral es *grande*.

TIP MI CONSEJO

Cuando $n < 30$, el teorema del límite central no garantiza que $\dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ sea aproximadamente normal.

Desafortunadamente, cuando el tamaño muestral n es pequeño, el estadístico $(\bar{x} - \mu)/(s/\sqrt{n})$ *no tiene* una distribución normal. Por tanto, todos los valores críticos de z que utilizamos en el capítulo 3 ya no son correctos. Por ejemplo, *no se puede decir* que se encontrará a no más de 1.96 errores estándar de μ 95% de las veces.

Este problema no es nuevo; fue estudiado por expertos en estadística y experimentadores a principios del siglo XX. Para hallar la distribución muestral de este estadístico, hay dos formas de proceder:

- Usar un método empírico. Saque repetidas muestras y calcule $(\bar{x} - \mu)/(s/\sqrt{n})$ para cada muestra. La distribución de frecuencia relativa que usted construya usando estos valores aproximarán la forma y ubicación de la distribución muestral.
- Usar un método matemático para deducir la función real de densidad o curva que describa la distribución muestral.

Este segundo método fue utilizado por un inglés llamado W. S. Gosset en 1908. Él dedujo una complicada fórmula para la función de densidad de

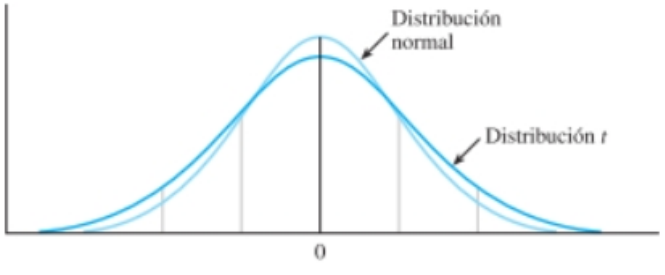
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

[†] Una prueba de muestra pequeña para el parámetro binomial p se presentará en el capítulo 11.

para muestras aleatorias de tamaño n desde una población normal y publicó sus resultados con el nombre de "Student". Desde entonces, la estadística se conoce como ***t* de Student** y tiene las siguientes características:

- Tiene forma de montículo y es simétrica alrededor de $t = 0$, igual que z.
- Es más variable que z, con "colas más pesadas"; esto es, la curva t no se aproxima al eje horizontal con la misma rapidez que z. Esto es porque el estadístico t abarca dos cantidades aleatorias, $\bar{x}$ y s, en tanto que el estadístico z tiene sólo la media muestral, $\bar{x}$. Se observa este fenómeno en la figura 4.1.
- La forma de la distribución t depende del tamaño muestral n. A medida que n aumenta, la variabilidad de t disminuye porque la estimación s de σ está basada en más y más información. En última instancia, cuando n sea infinitamente grande, ¡las distribuciones t y z son idénticas!

FIGURA 4.1
z normal estándar y la distribución t con 5 grados de libertad



El divisor $(n - 1)$ en la fórmula para la varianza muestral s^2 se denomina **número de grados de libertad (*gl*) asociados con s^2**. Determina la *forma* de la distribución t. El origen del término *grados de libertad* es teórico y se refiere al número de desviaciones independientes elevadas al cuadrado en s^2 existentes para estimar σ^2. Estos grados de libertad pueden cambiar para diferentes aplicaciones, y como especifican la distribución t correcta que se debe usar es necesario recordar que hay que calcular los grados de libertad correctos para cada aplicación.

TIP MI CONSEJO

Para una t de una muestra, $df = n - 1$

La tabla de probabilidades para la distribución z normal estándar ya no es útil para calcular valores críticos o valores p para el estadístico t. En lugar de ello, se usará la tabla 4 del Apéndice que se reproduce parcialmente en la tabla 4.1. Al indizar un número particular de grados de libertad, la tabla registra t_a, un valor de t que tiene área a de cola a su derecha, como se muestra en la figura 4.2.

FIGURA 4.2
Valores tabulados de la t de Student

TABLA 4.1 **Formato de la tabla *t* de Student tomado de la tabla 4 del Apéndice**

gl	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	*df*
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
inf.	1.282	1.645	1.960	2.326	2.576	inf.

APPLET EN LÍNEA

Probabilidades *t* de Student comparando *t* y *z*

EJEMPLO 4.1

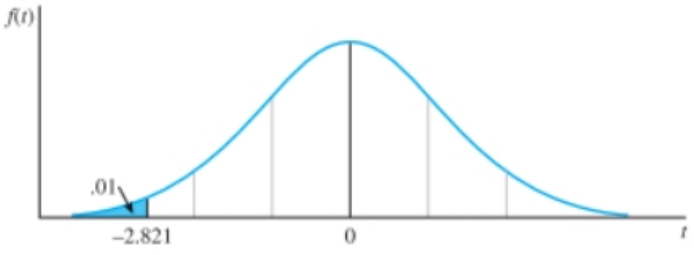
Para una distribución *t* con 5 grados de libertad, el valor de *t* que tiene área .05 a su derecha se encuentra en la fila 5 en la columna marcada $t_{.050}$. Para esta distribución *t* particular, el área a la derecha de $t = 2.015$ es .05; sólo 5% de todos los valores del estadístico *t* rebasarán este valor.

EJEMPLO 4.2

Suponga que usted tiene una muestra de tamaño $n = 10$ de una distribución normal. Encuentre un valor de *t* tal que sólo 1% de todos los valores de *t* sea más pequeño.

Solución Los grados de libertad que especifican la distribución *t* correcta son $gl = n - 1 = 9$, y el valor *t* necesario debe estar en la parte inferior de la distribución, con área .01 a su izquierda, como se ve en la figura 4.3. Como la distribución *t* es simétrica alrededor de 0, este valor es simplemente el negativo del valor en el lado derecho con área .01 a su derecha, o $-t_{.01} = -2.821$.

FIGURA 4.3
Distribución *t* para el ejemplo 4.2



Podría preguntarse por qué los grados de libertad (*gl*) en la tabla 4.1 saltan de $gl = 29$ a $gl = inf.$ (infinito). Los valores críticos de *t* para varios grados de libertad entre 29 y 300 se dan en la figura 4.4. Notará que el valor de *t* para la misma área de la cola derecha disminuye conforme aumentan los grados de libertad. Cuando los grados de libertad se vuelven infinitamente grandes (*inf.*), el valor de *t* es igual al valor de *z*, el cual se proporciona en la última fila de la figura 4.4.

FIGURA 4.4

Valores críticos de la *t* de Student para grados de libertad entre *gl* = 29 y *gl* = *infinito*

	Área de la cola derecha		
gl	0.05	0.025	0.01
29	1.699	2.045	2.462
49	1.677	2.010	2.405
69	1.667	1.995	2.382
100	1.660	1.984	2.364
200	1.653	1.972	2.345
300	1.650	1.968	2.339
inf.	1.645	1.96	2.326

Al mismo tiempo, conforme aumentan los grados de libertad, la forma de la distribución *t* se vuelve menos variable hasta que al final se ve como (y es) la distribución normal estándar. Note que cuando los grados de libertad con *t* son $gl = 300$, casi no hay diferencia. Cuando $gl = 29$ y $n = 30$ los valores críticos de *t* son bastante cercanos a sus contrapartes normales; esto puede explicar por qué se usa la línea divisoria arbitraria entre una muestra grande y pequeña como $n = 30$. En lugar de producir una tabla *t* con muchos más valores críticos, los valores críticos de *z* son suficientes cuando *n* llega a 30.

Suposiciones detrás de la distribución *t* de Student

Los valores críticos de *t* le permiten hacer inferencias confiables *sólo si* usted sigue todas las reglas; esto es, su muestra debe cumplir estos requisitos especificados por la distribución *t*:

- La muestra debe ser seleccionada al azar.
- La población de la que haga muestreo debe estar normalmente distribuida.

TIP MI CONSEJO

Suposiciones para *t* de una muestra:

- Muestra aleatoria
- Distribución normal

Estos requisitos parecen bastante restrictivos. ¿Cómo podría conocer la forma de la distribución de probabilidad para toda la población si nada más tiene una muestra? Pero si éste fuera un problema serio el estadístico *t* podría usarse sólo en situaciones muy limitadas. Por fortuna, la forma de la distribución *t* no es afectada en mucho mientras la población muestreada tenga una distribución *aproximadamente en forma de montículo*. Los expertos dicen que el estadístico *t* es **robusto**, lo cual significa que su distribución no cambia de manera significativa cuando se viola la suposición de normalidad.

¿Cómo se puede saber si la muestra es de una población normal? Aun cuando hay procedimientos estadísticos diseñados para este fin, la forma más fácil y rápida de verificar la normalidad es usar las técnicas gráficas del capítulo 2: trazar una gráfica de puntos o construir una gráfica de tallo y hoja. Mientras esta gráfica tienda a "hacerse montículo" en el centro, se puede estar razonablemente seguro al usar el estadístico *t* para hacer inferencias.

El requisito de muestreo aleatorio, por otra parte, es bastante crítico si se desea producir inferencias confiables. Si la muestra no es aleatoria, o si no se *comporta al menos como* muestra aleatoria, entonces los resultados de su muestra pueden ser afectados por algún factor desconocido y las conclusiones pueden ser incorrectas. Cuando diseñe un experimento o lea acerca de experimentos realizados por otros investigadores, ¡observe de manera crítica la forma en que los datos han sido recolectados!

4.2 INFERENCIA CON MUESTRAS PEQUEÑAS PARA LA MEDIA DE LA POBLACIÓN

Al igual que con una inferencia de muestra grande, una inferencia de muestra pequeña puede comprender ya sea **estimación** o **prueba de hipótesis**, dependiendo de la preferencia del experimentador. En capítulos previos explicamos los aspectos básicos de estos dos tipos de inferencia y los usamos de nuevo ahora con una estadística muestral diferente, $t = (\bar{x} - \mu)/(s/\sqrt{n})$, y una distribución de muestreo diferente, la *t* de Student, con $(n - 1)$ grados de libertad.

PRUEBA DE HIPÓTESIS DE MUESTRA PEQUEÑA PARA μ

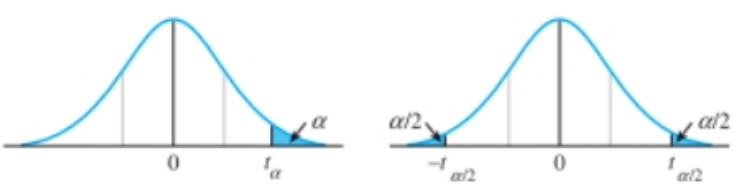
1. Hipótesis nula: H_0: $\mu = \mu_0$
2. Hipótesis alternativa:

Prueba de una cola	Prueba de dos colas
H_a: $\mu > \mu_0$ (o, H_a: $\mu < \mu_0$)	H_a: $\mu \neq \mu_0$

3. Estadístico de prueba: $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$
4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	Prueba de dos colas
$t > t_\alpha$ (o $t < -t_\alpha$ cuando la hipótesis alternativa es H_a: $\mu < \mu_0$)	$t > t_{\alpha/2}$ o $t < -t_{\alpha/2}$

o cuando el valor $p < \alpha$



Los valores críticos de t, t_α y $t_{\alpha/2}$ están basados en $(n - 1)$ grados de libertad. Estos valores tabulados se encuentran usando la tabla 4 del Apéndice.

Suposición: La muestra se selecciona al azar de una población normalmente distribuida.

INTERVALO DE CONFIANZA DE MUESTRA PEQUEÑA $(1 - \alpha)100\%$ PARA μ

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$$

donde $s/\sqrt{n}$ es el error estándar estimado de $\bar{x}$, a veces conocido como **error estándar de la media**.

EJEMPLO 4.3

Un nuevo proceso para producir diamantes sintéticos puede ser operado a un nivel rentable sólo si el promedio del peso de éstos es mayor que .5 quilates. Para evaluar la rentabilidad del proceso se generan seis diamantes que registran pesos de .46, .61, .52, .48, .57 y .54 quilates. ¿Estas seis mediciones presentan suficiente evidencia para indicar que el promedio del peso de los diamantes producidos por el proceso es más de.5 quilates?

Solución La población de pesos de diamantes producidos por este nuevo proceso tiene media μ, y usted puede empezar la prueba formal de hipótesis en pasos:

1–2 **Hipótesis nula y alternativa:**

H_0: $\mu = .5$ contra H_a: $\mu > .5$

3 **Estadístico de prueba**: Usted puede usar su calculadora para verificar que la media y desviación estándar para los pesos de los seis diamantes son .53 y .0559, respectivamente. El estadístico de prueba es un estadístico t, calculado como

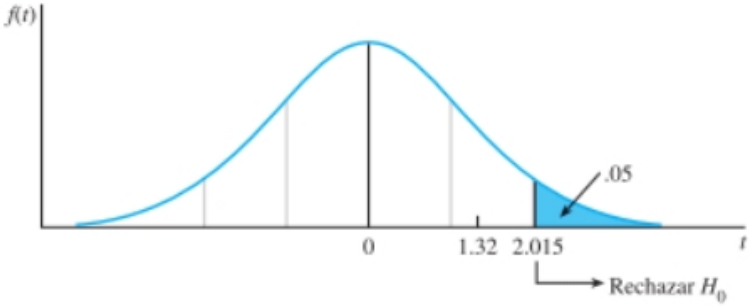
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{.53 - .5}{.0559/\sqrt{6}} = 1.32$$

Al igual que con las pruebas de muestra grande, el estadístico de prueba proporciona evidencia para rechazar o aceptar H_0 dependiendo de qué tan lejos se encuentre del centro de la distribución t.

4 **Región de rechazo**: Si elige un nivel de significancia de 5% ($\alpha = .05$), la región de rechazo de la cola derecha se encuentra usando los valores críticos de t de la tabla 4 del Apéndice. Con $gl = n - 1 = 5$, se puede rechazar H_0 si $t > t_{.05} = 2.015$, como lo muestra la figura 4.5.

5 **Conclusión**: Como el valor calculado del estadístico de prueba, 1.32, no cae en la región de rechazo, no se puede rechazar H_0. Los datos no presentan suficiente evidencia para indicar que el peso medio de los diamantes exceda de .5 quilates.

FIGURA 4.5
Región de rechazo para el ejemplo 4.3



TIP MI CONSEJO

Un intervalo de confianza de 95% nos dice que, si construyéramos muchos de estos intervalos (todos los cuales tendrían puntos extremos ligeramente diferentes), 95% de ellos se acercarían a la media poblacional

La conclusión para *aceptar* H_0 requeriría el difícil cálculo de β, la probabilidad de un error tipo II. Para evitar este problema, elegimos *no rechazar* H_0. Podemos entonces calcular el límite inferior para μ usando un límite inferior de confianza de un lado y muestra pequeña. Este límite es similar al límite de confianza de un lado y muestra grande, excepto que la z_α crítica es sustituida por una t_α crítica de la tabla 4. Para este ejemplo, un límite inferior de confianza de un lado para μ es:

$$\bar{x} - t_\alpha \frac{s}{\sqrt{n}}$$

$$.53 - 2.015 \frac{.0559}{\sqrt{6}}$$

$$.53 - .046$$

El límite inferior de 95% para μ es $\mu > .484$. El rango de posibles valores incluye la media de los pesos de los diamantes tanto menores como mayores que .5; esto confirma el fracaso de nuestra prueba para demostrar que μ excede de .5.

Hay dos formas de efectuar una prueba de hipótesis:

- **El método del valor crítico**: Proponga una región de rechazo basada en los valores críticos de la distribución muestral del estadístico. Si el estadístico de prueba cae en la región de rechazo, se puede rechazar H_0.
- **El método del valor *p***: Calcule el valor p con base en el valor observado del estadístico de prueba. Si el valor p es menor que el nivel de significancia, α, se puede rechazar H_0. Si no hay nivel de significancia *preestablecido*, use las normas de la sección 9.3 para juzgar la significancia estadística de sus resultados muestrales.

Usamos el primer método en la solución del ejemplo 4.3 y el segundo método para resolver el ejemplo 4.4.

EJEMPLO 4.4

Las etiquetas en latas de 1 galón de pintura por lo general indican el tiempo de secado y el área que puede cubrir una capa. Casi todas las marcas de pintura indican que, en una capa, 1 galón cubrirá entre 250 y 500 pies cuadrados, dependiendo de la textura de la superficie a pintarse. Un fabricante, sin embargo, afirma que 1 galón de su pintura cubrirá 400 pies cuadrados de área superficial. Para probar su afirmación, una muestra aleatoria de 10 latas de 1 galón de pintura blanca se empleó para pintar 10 áreas idénticas usando la misma clase de equipo. Las áreas reales (en pies cuadrados) cubiertas por estos 10 galones de pintura se dan a continuación:

310	311	412	368	447
376	303	410	365	350

¿Los datos presentan suficiente evidencia para indicar que el promedio de la cobertura difiere de 400 pies cuadrados? Encuentre el valor p para la prueba y úselo para evaluar la significancia estadística de los resultados.

Solución Para probar lo dicho, las hipótesis que deben probarse son

$$H_0: \mu = 400 \quad \text{contra} \quad H_a: \mu \neq 400$$

La media muestral y desviación estándar para los datos registrados son

$$\bar{x} = 365.2 \qquad s = 48.417$$

y el estadístico de prueba es

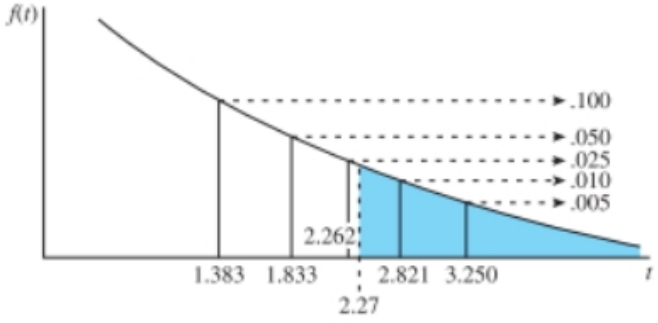
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{365.2 - 400}{48.417/\sqrt{10}} = -2.27$$

El valor p para esta prueba es la probabilidad de observar un valor de la estadística t como contradictorio a la hipótesis nula como el observado para este conjunto de datos, es decir, $t = -2.27$. Como ésta es una prueba de dos colas, el valor p es la probabilidad de que $t \leq -2.27$ o $t \geq 2.27$.

A diferencia de la tabla z, la tabla para t da los valores de t correspondientes a áreas de cola superior iguales a .100, .050, .025, .010 y .005. En consecuencia, sólo puede aproximarla al área de cola superior que corresponda a la probabilidad de que $t > 2.27$. Como el estadístico t para esta prueba está basado en 9 gl, nos referimos a la fila correspondiente a $gl = 9$ de la tabla 4. Los cinco valores críticos para varias áreas de cola se muestran en la figura 4.6, un agrandamiento de la cola de la distribución t con 9 grados de libertad. El valor $t = 2.27$ cae entre $t_{.025} = 2.262$ y $t_{.010} = 2.821$. Por tanto, el área de cola derecha correspondiente a la probabilidad de que $t > 2.27$ está entre .01 y .025. Como esta área representa sólo la mitad del valor p, se escribe

$$.01 < \frac{1}{2}(\text{valor } p) < .025 \quad \text{o} \quad .02 < \text{valor } p < .05$$

FIGURA 4.6
Cálculo del valor p para el ejemplo 4.4 (área sombreada $\frac{1}{2}$ = valor p)



¿Qué nos dice esto acerca de la significancia de los resultados estadísticos? Para rechazar H_0, el valor p debe ser menor que el nivel de significancia especificado, α. En consecuencia, podría rechazarse H_0 al nivel de 5% pero no al nivel de 2% o 1%. Por tanto, el valor p para esta prueba por lo general sería informado por el experimentador como

$$\text{valor } p < .05 \text{ (o a veces } P < .05)$$

Para esta prueba de hipótesis, H_0 es rechazada al nivel de significancia de 5%. Hay suficiente evidencia para indicar que el promedio de la cobertura difiere de 400 pies cuadrados.

APPLET EN LÍNEA
Prueba de muestra pequeña de una media poblacional

¿Dentro de qué límites *realmente* cae este promedio de cobertura? Un intervalo de confianza de 95% da los límites superior e inferior para μ como

$$\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$365.2 \pm 2.262\left(\frac{48.417}{\sqrt{10}}\right)$$

$$365.2 \pm 34.63$$

Por tanto, usted puede estimar que el promedio de área cubierta por 1 galón de esta marca de pintura está en el intervalo 330.6 a 399.8. Una estimación más precisa del intervalo (un intervalo más corto) puede obtenerse por lo general al aumentar el tamaño muestral. Observe que el límite superior de este intervalo es muy cercano al valor de 400 pies cuadrados, que es la cobertura marcada en la leyenda de la lata. Esto coincide con el hecho de que el valor observado de $t = -2.27$ es sólo ligeramente menor que el valor crítico de cola izquierda de $t_{.025} = -2.262$, haciendo que el valor p sea sólo ligeramente menor que .05.

Muchos paquetes de estadística para computadoras contienen programas que ponen en práctica la prueba t de Student o construyen un intervalo de confianza para μ. Aunque *MS Excel* no tiene un comando único que implemente estos procedimientos, usted puede usar la herramienta de función en *Excel* para encontrar por sí mismo el estadístico de prueba, el valor p, y los límites de confianza superior e inferior. *MINITAB*, sin embargo, calcula y reporta todos estos valores con un conjunto de comandos, permitiéndole obtener con rapidez y precisión conclusiones sobre la significancia estadística de los resultados. Los resultados del *MINITAB* de prueba t de una muestra y de procedimientos de intervalo de confianza se dan en la figura 10.7. Además del valor observado de $t = -2.27$ y el intervalo de confianza (330.6, 399.8), la salida da la media muestral, la desviación muestral estándar, el error estándar de la media (media de error estándar $= s/\sqrt{n}$), y el valor p exacto de la prueba ($P = .049$). Esto es consistente con el rango para el valor p que encontramos usando la tabla 4 del Apéndice:

$$.02 < \text{valor } p < .05$$

Encontrará instrucciones para generar esta salida *MINITAB* en la sección 4.7.

FIGURA 4.7
Salida impresa *MINITAB* para el ejemplo 4.4

T de una muestra: área

```
Test of mu = 400 vs not = 400

Variable        N     Mean       StDev        SE Mean
Area           10    365.2        48.4           15.3

Variable        95% CI                   T         P
Area       (330.6, 399.8)            -2.27     0.049
```

Se puede ver el valor de usar la salida impresa de computadora para evaluar resultados estadísticos:

- El valor *p* exacto elimina la necesidad de tablas y valores críticos.
- Todos los cálculos numéricos son realizados por el usuario.

El trabajo más importante, que se deja al experimentador, es *interpretar* los resultados en términos de su ¡significancia práctica!

4.3 INFERENCIA CON MUESTRAS PEQUEÑAS PARA LA DIFERENCIA ENTRE DOS MEDIAS

El escenario físico para el problema considerado en esta sección es el mismo que el de la sección 3.3, excepto que los tamaños muestrales ya no son grandes. Se extraen muestras aleatorias independientes de n_1 y n_2 mediciones de dos poblaciones, con medias y varianzas μ_1, σ_1^2, μ_2 y σ_2^2, y el objetivo del experimentador es hacer inferencias acerca de $(\mu_1 - \mu_2)$, la diferencia entre las dos medias poblacionales.

Cuando los tamaños muestrales son pequeños ya no se puede confiar en el teorema del límite central para asegurar que las medias muestrales sean normales, pero, si las poblaciones originales *son normales*, entonces la distribución muestral de la diferencia en las medias muestrales, $(\bar{x}_1 - \bar{x}_2)$, será normal (incluso para muestras pequeñas) con media $(\mu_1 - \mu_2)$ y error estándar

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

TIP MI CONSEJO

Suposiciones para la prueba *t* de dos muestras (independientes):

- Muestras aleatorias independientes
- Distribuciones normales
- $\sigma_1 = \sigma_2$

En los capítulos 1 y 3 usted usó las varianzas muestrales, s_1^2 y s_2^2, para calcular una *estimación* del error estándar, que entonces se utilizó para formar un intervalo de confianza de muestra grande o una prueba de hipótesis basada en el estadístico *z* de muestras grandes:

$$z \approx \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Desafortunadamente, cuando los tamaños muestrales son pequeños, este estadístico no tiene una distribución aproximadamente normal ni tiene una distribución *t* de Student. Para formar una estadística con una distribución de muestreo que pueda deducirse en forma teórica es necesario hacer una suposición más.

Suponga que la variabilidad de la medición en las dos poblaciones normales es la misma y puede ser medida por una varianza común σ^2. Esto es, *ambas poblaciones tienen exactamente la misma forma*, y $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Entonces el error estándar de la diferencia en las dos medias muestrales es

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Se puede demostrar matemáticamente que, si se usa la estimación muestral apropiada s^2 para la varianza poblacional σ^2, entonces el estadístico de prueba resultante,

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

tiene una *distribución t de Student*. El único problema restante es hallar la estimación muestral s^2 y el número apropiado de *grados de libertad* para el estadístico t.

Recuerde que la varianza poblacional σ^2 describe la forma de las distribuciones normales de donde provienen las muestras del experimentador, de modo que s_1^2 o s_2^2 le darían una estimación de σ^2. Pero, ¿por qué usar sólo una cuando ambas dan información? Un mejor procedimiento es combinar la información en ambas varianzas muestrales usando un *promedio ponderado*, en el que los pesos están determinados por la cantidad relativa de información (el número de mediciones) en cada muestra. Por ejemplo, si la primera muestra contenía el doble de mediciones que la segunda, se podría considerar dar a la primera varianza muestral el doble de peso. Para obtener este resultado, use esta fórmula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Recuerde de la sección 4.2 que los grados de libertad para el estadístico t de una muestra son $(n - 1)$, el denominador de la estimación muestral s^2. Como s_1^2 tiene $(n_1 - 1)$ *gl* y s_2^2 tiene $(n_2 - 1)$ *gl*, el número total de grados de libertad es la suma

$$(n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

mostrada en el denominador de la fórmula para s^2.

CÁLCULO DE s^2

- Si dispone de una calculadora científica, calcule cada una de las dos desviaciones estándar muestrales s_1 y s_2 por separado, usando el procedimiento de entrada de datos para su calculadora particular. Estos valores son elevados al cuadrado y se usan en esta fórmula:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

TIP MI CONSEJO

Para la prueba *t* de dos muestras (independientes) $df = n_1 + n_2 - 2$

Se puede demostrar que s^2 es un estimador no sesgado de la varianza poblacional común σ^2. Si s^2 se usa para estimar σ^2 y si las muestras se han sacado al azar e independientemente de poblaciones normales con una varianza común, entonces el estadístico

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

tiene una distribución t de Student con $(n_1 + n_2 - 2)$ grados de libertad. La estimación de muestra pequeña y procedimientos de prueba para la diferencia entre dos medias se dan a continuación.

PRUEBA DE HIPÓTESIS RESPECTO A LA DIFERENCIA ENTRE DOS MEDIAS: MUESTRAS ALEATORIAS INDEPENDIENTES

1. La hipótesis nula: H_0: $(\mu_1 - \mu_2) = D_0$, donde D_0 es alguna diferencia especificada que el experimentador desea probar. Para numerosas pruebas, el experimentador hará una hipótesis de que no hay diferencia entre μ_1 y μ_2; esto es, $D_0 = 0$.
2. Hipótesis alternativa:

Prueba de una cola	**Prueba de dos colas**
H_a: $(\mu_1 - \mu_2) > D_0$	H_a: $(\mu_1 - \mu_2) \neq D_0$
[o H_a: $(\mu_1 - \mu_2) < D_0$]	

3. Estadístico de prueba: $$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$ donde

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

4. Región de rechazo: rechace H_0 cuando:

Prueba de una cola	**Prueba de dos colas**
$t > t_\alpha$	$t > t_{\alpha/2}$ o $t < -t_{\alpha/2}$
[o $t < -t_\alpha$ cuando la hipótesis alternativa es H_a: $(\mu_1 - \mu_2) < D_0$]	
o cuando el valor $p < \alpha$	

Los valores críticos de t, t_α y $t_{\alpha/2}$ están basados en $(n_1 + n_2 - 2)$ *df*. Los valores tabulados se pueden hallar usando la tabla 4 del Apéndice.

Suposiciones: Las muestras se seleccionan al azar y en forma independiente de poblaciones distribuidas normalmente. Las varianzas de las poblaciones σ_1^2 y σ_2^2 son iguales.

INTERVALO DE CONFIANZA $(1 - \alpha)100\%$ DE MUESTRA PEQUEÑA PARA $(\mu_1 - \mu_2)$ CON BASE EN MUESTRAS ALEATORIAS INDEPENDIENTES

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

donde s^2 es la estimación agrupada de σ^2.

EJEMPLO 4.5

Puede tomarse un curso con crédito ya sea asistiendo a clases en horas y días fijos, o en sesiones en línea que el estudiante puede tomar a su propio paso y en los tiempos que elija. El coordinador del curso desea determinar si estas dos formas de tomar el curso resultaron en una diferencia significativa en rendimiento medido por el examen final. La tabla 4.2 da las calificaciones en un examen con 45 puntos posibles para un grupo de $n_1 = 9$ estudiantes que tomaron el curso en línea y un segundo grupo de $n_2 = 9$ estudiantes que lo tomaron en clases

convencionales. ¿Estos datos presentan suficiente evidencia para indicar que el promedio de las calificaciones de los estudiantes que tomaron el curso en línea es significativamente más alto que el de quienes asistieron a una clase convencional?

TABLA 4.2 **Calificaciones de examen para presentaciones en línea y en salón de clase**

En línea	Salón de clases
32	35
37	31
35	29
28	25
41	34
44	40
35	27
31	32
34	31

Solución Sean μ_1 y μ_2 la media de las calificaciones para el grupo en línea y el grupo del salón de clase, respectivamente. Por tanto, como se busca evidencia para apoyar la teoría de que $\mu_1 > \mu_2$, se puede probar la hipótesis nula

$$H_0 : \mu_1 = \mu_2 \qquad [\text{o } H_0 : (\mu_1 - \mu_2) = 0]$$

contra la hipótesis alternativa

$$H_a : \mu_1 > \mu_2 \qquad [\text{o } H_a : (\mu_1 - \mu_2) > 0]$$

Para efectuar la prueba t para estas dos muestras independientes se debe suponer que las poblaciones muestreadas son normales y tienen la misma varianza σ^2. ¿Es esto razonable? Las gráficas de tallo y hoja de la figura 4.8 muestran al menos un patrón de "montículo", de modo que la suposición de normalidad es razonable.

FIGURA 4.8
Gráficas de tallo y hoja para el ejemplo 4.5

En línea		Salón de clases	
2	8	2	579
3	124	3	1124
3	557	3	5
4	14	4	0

TIP MI CONSEJO

Las gráficas de tallo y hoja pueden ayudar a decidir si la suposición de normalidad es razonable

Además, las desviaciones estándar de las dos muestras, calculadas como

$$s_1 = 4.9441 \quad \text{y} \quad s_2 = 4.4752$$

no son tan diferentes para que dudemos de que las dos distribuciones tengan la misma forma. Si se hacen estas dos suposiciones y se calcula (usando precisión total) la estimación agrupada de la varianza común como

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{8(4.9441)^2 + 8(4.4752)^2}{9 + 9 - 2} = 22.2361$$

se puede calcular entonces el estadístico de prueba,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{35.22 - 31.56}{\sqrt{22.2361\left(\frac{1}{9} + \frac{1}{9}\right)}} = 1.65$$

La hipótesis alternativa $H_a : \mu_1 > \mu_2$ o, lo que es equivalente, $H_a : (\mu_1 - \mu_2) > 0$ implica que el experimentador debe usar una prueba de una cola en la cola superior de la distribución t con

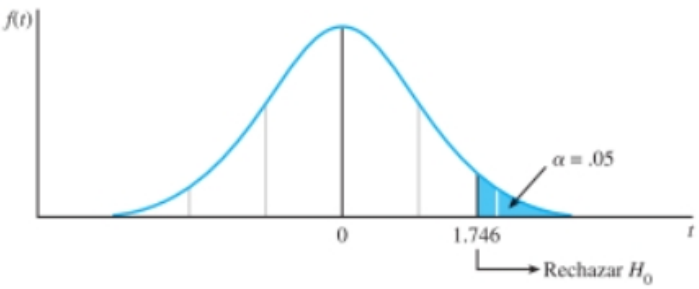
TIP MI CONSEJO

Si usa calculadora, ¡no redondee sino hasta el paso final!

$(n_1 + n_2 - 2) = 16$ grados de libertad. Se puede hallar el valor crítico apropiado para una región de rechazo con $\alpha = .05$ en la tabla 4 del Apéndice, y H_0 será rechazada si $t > 1.746$. Comparando el valor observado del estadístico de prueba $t = 1.65$ con el valor crítico $t_{.05} = 1.746$, no se puede rechazar la hipótesis nula (véase la figura 4.9). Hay insuficiente evidencia para indicar que el promedio de las calificaciones del curso en línea sea más alto que el promedio de las calificaciones del curso convencional al nivel de significancia de 5%.

FIGURA 4.9

Región de rechazo para el ejemplo 4.5



EJEMPLO 4.6

Encuentre el valor p que sería reportado para la prueba estadística del ejemplo 4.5.

Solución El valor observado de t para esta prueba de una cola es $t = 1.65$. Por tanto,

$$\text{valor } p = P(t > 1.65)$$

MI APPLET EN LÍNEA

Prueba t de dos muestras: muestras independientes

para un estadístico t con 16 grados de libertad. Recuerde que no se puede obtener esta probabilidad directamente de la tabla 4 del Apéndice; sólo se puede *limitar* el valor p usando los valores críticos de la tabla. Como el valor observado, $t = 1.65$, está entre $t_{.100} = 1.337$ y $t_{.050} = 1.746$, el área de cola a la derecha de 1.65 está entre .05 y .10. El valor p para esta prueba se informaría como

$$.05 < \text{valor } p < .10$$

Como el valor p es mayor que .05, casi todos los investigadores informarían los resultados como *no significativos*.

EJEMPLO 4.7

Use un límite inferior de confianza de 95% para estimar la diferencia $(\mu_1 - \mu_2)$ en el ejemplo 4.5. ¿El límite de confianza inferior indica que el promedio en línea es significativamente más alto que el promedio de salón de clase?

Solución El límite inferior de confianza toma una forma familiar, es decir, el estimador puntual $(\bar{x}_1 - \bar{x}_2)$ menos una cantidad igual a t_α veces del error estándar del estimador. Sustituyendo en la fórmula, se puede calcular el límite inferior de confianza de 95%:

$$(\bar{x}_1 - \bar{x}_2) - t_\alpha \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$(35.22 - 31.56) - 1.746\sqrt{22.2361\left(\frac{1}{9} + \frac{1}{9}\right)}$$

$$3.66 - 3.88$$

o $(\mu_1 - \mu_2) > -.22$. Como el valor $(\mu_1 - \mu_2) = 0$ está incluido en el intervalo de confianza, es posible que las dos medias sean iguales. Hay insuficiente evidencia para indicar que el promedio en línea sea más alto que el promedio en salón de clase.

TIP MI CONSEJO

Mayor s^2/menor $s^2 < 3$

$\Leftrightarrow$

suposición de varianza es razonable

El procedimiento de dos muestras que usa una estimación agrupada de la varianza común σ^2 se apoya en cuatro importantes suposiciones:

- Las muestras deben ser *seleccionadas al azar*. Las muestras no seleccionadas al azar pueden introducir sesgo en el experimento y así alterar los niveles de significancia que el experimentador informe.
- Las muestras deben ser *independientes*. Si no es así, éste no es el procedimiento estadístico apropiado. En la sección 4.5 analizamos otro procedimiento para muestras dependientes.
- Las poblaciones de las cuales se muestrea deben ser *normales*. No obstante, las desviaciones moderadas desde la normalidad no afectan seriamente la distribución del estadístico de prueba, en especial si los tamaños muestrales son casi iguales.
- Las *varianzas poblacionales deben ser iguales* o casi iguales para asegurar que los procedimientos sean válidos.

Si las varianzas poblacionales están lejos de ser iguales, hay un procedimiento alternativo para estimación y prueba que tiene una distribución *t aproximada* en muestreo repetido. Como regla práctica, se debe usar este procedimiento si la razón entre las dos varianzas muestrales,

$$\frac{\text{Mayor } s^2}{\text{Menor } s^2} > 3$$

Como las varianzas poblacionales no son iguales, el estimador agrupado s^2 ya no es apropiado, y cada varianza poblacional debe ser estimada por su correspondiente varianza muestral. El estadístico de prueba resultante es

$$\frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Cuando los tamaños muestrales sean *pequeños*, los valores críticos para este estadístico se encuentran usando grados de libertad aproximados por la fórmula

$$gl \approx \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{(n_1 - 1)} + \frac{(s_2^2/n_2)^2}{(n_2 - 1)}}$$

Los grados de libertad se toman como la parte entera de este resultado.

Se pueden usar paquetes computarizados como el *MINITAB* y el *MS Excel* para poner en práctica este procedimiento, a veces llamado *aproximación de Satterthwaite*, así como el *método agrupado* descrito antes. De hecho, algunos experimentadores eligen analizar sus datos usando *ambos* métodos. Mientras los dos análisis lleven a las mismas conclusiones, no es necesario preocuparse con la igualdad o desigualdad de varianzas.

Las salidas *MINITAB* y *Excel*, resultantes del método agrupado de análisis para los datos del ejemplo 4.5, se muestran en las figuras 10.10a y b. Observe que la razón entre las dos varianzas muestrales, $(4.94/4.48)^2 = 1.22$ es menor que 3, lo cual hace aproximado el método agrupado. El valor calculado de $t = 1.65$ y el valor p exacto $= .059$ con 16 grados de libertad se muestran en ambas salidas. El valor p exacto hace muy fácil determinar la significancia o no significancia de los resultados muestrales. Se encontrarán instrucciones para generar esta salida en la sección 4.7.

FIGURA 4.10a
Salida de *MINITAB* para el ejemplo 4.5

Prueba t de dos muestras y CI: en línea, salón de clase

```
Two-sample T for Online vs Classroom
             N      Mean      StDev   SE Mean
Online       9     35.22       4.94       1.6
Classroom    9     31.56       4.48       1.5

Difference = mu (Online) - mu (Classroom)
Estimate for difference: 3.67
95% lower bound for difference: -0.21
T-Test of difference = 0 (vs >): T-Value = 1.65  P-Value = 0.059  DF = 16
Both use Pooled StDev = 4.7155
```

FIGURA 4.10b
Salida de Excel para el ejemplo 4.5

D	E	F
t-Test: Two-Sample Assuming Equal Variances		
	Online	*Classroom*
Mean	35.222	31.556
Variance	24.444	20.028
Observations	9	9
Pooled Variance	22.236	
Hypothesized Mean Difference	0	
df	16	
t Stat	1.649	
P(T<=t) one-tail	0.059	
t Critical one-tail	1.746	
P(T<=t) two-tail	0.119	
t Critical two-tail	2.120	

Si hay razón para pensar que las suposiciones de normalidad han sido violadas se puede probar un cambio en la ubicación de dos distribuciones poblacionales usando la prueba de suma de rango no paramétrica de Wilcoxon del capítulo 11. Este procedimiento de prueba, que requiere menos suposiciones respecto a la naturaleza de las distribuciones de probabilidad poblacionales, es casi igualmente sensible para detectar una diferencia en medias poblacionales cuando las condiciones necesarias para la prueba t se satisfagan. Puede ser más sensible cuando no se satisfaga la suposición de normalidad.

4.4 PRUEBA PARA DIFERENCIAS PAREADAS

Para comparar las cualidades de desgaste de dos tipos de llantas de automóvil, A y B, una llanta de tipo A y una de tipo B se asignaron al azar y se montaron en las ruedas traseras a cada uno de cinco automóviles. Éstos se hicieron correr un número especificado de millas y se registró la cantidad de desgaste para cada llanta. Estas mediciones aparecen en la tabla 4.3. ¿Los datos presentan suficiente evidencia para indicar una diferencia en el promedio de desgaste para los dos tipos de llantas?

TABLA 4.3 **Promedio de desgaste para dos tipos de llantas**

Automóvil	Llanta A	Llanta B
1	10.6	10.2
2	9.8	9.4
3	12.3	11.8
4	9.7	9.1
5	8.8	8.3
	$\bar{x}_1 = 10.24$	$\bar{x}_2 = 9.76$
	$s_1 = 1.316$	$s_2 = 1.328$

La tabla 4.3 muestra una diferencia de $(\bar{x}_1 - \bar{x}_2) = (10.24 - 9.76) = .48$ entre las dos medias muestrales, en tanto que las desviaciones estándar de ambas muestras son aproximadamente 1.3. Dada la variabilidad de los datos y el pequeño número de mediciones, ésta es una diferencia bastante pequeña y es probable que usted no sospeche una diferencia en el promedio de desgaste para los dos tipos de llantas. Revisemos las sospechas usando los métodos de la sección 4.3.

Vea el análisis *MINITAB* de la figura 4.11. La prueba *t agrupada* de dos muestras se usa para probar la diferencia en las medias basada en dos muestras aleatorias independientes. El valor calculado de t usado para probar la hipótesis nula $H_0 : \mu_1 = \mu_2$ es $t = .57$ con valor $p = .582$, valor que no es casi suficientemente pequeño para indicar una diferencia significativa en las dos medias poblacionales. El correspondiente intervalo de confianza de 95%, dado como

$$-1.448 < (\mu_1 - \mu_2) < 2.408$$

es bastante ancho y no indica una diferencia significativa en las medias poblacionales.

FIGURA 4.11

Salida impresa *MINITAB* usando prueba *t* para muestras independientes para los datos de llantas

Prueba T de dos muestras y CI: llanta A, llanta B

```
Two-sample T for Tire A vs Tire B
        N       Mean       StDev    SE Mean
Tire A  5      10.24        1.32        0.59
Tire B  5       9.76        1.33        0.59

Difference = mu (Tire A) - mu (Tire B)
Estimate for difference: 0.480
95% CI for difference: (-1.448, 2.408)
T-Test of difference = 0 (vs not =): T-Value = 0.57  P-Value = 0.582  DF = 8
Both use Pooled StDev = 1.3221
```

Analice de nuevo los datos y observará que la medición del desgaste para el tipo A es mayor que el correspondiente valor para el tipo B, para *cada uno* de los cinco automóviles. ¿No sería esto improbable, si en realidad no hay diferencia entre los dos tipos de llantas?

Considere una prueba intuitiva sencilla, basada en la distribución binomial. Si no hay diferencia en el desgaste medio para los dos tipos de llantas, entonces es igualmente probable o no probable que la llanta A muestre más desgaste que la llanta B. Los cinco automóviles entonces corresponden a cinco intentos binomiales con $p = P$ (llanta A muestra más desgaste que la llanta B) $= .5$.

¿Es poco común el valor observado de $x = 5$ diferencias positivas mostradas en la tabla 4.4? La probabilidad de observar $x = 5$ o el igualmente probable valor de $x = 0$ se puede hallar en la tabla 1 del Apéndice que son $2(.031) = .062$, que es bastante pequeño en comparación con la probabilidad de la más potente prueba t, que tenía un valor p de .58. ¿No es peculiar que la prueba t, que usa más información (la muestra real de mediciones) que la prueba binomial, no dé suficiente información para rechazar la hipótesis nula?

Hay una explicación para esta inconsistencia. La prueba t descrita en la sección 4.3 *no* es el estadístico de prueba apropiado para aplicarlo a nuestro ejemplo. El procedimiento de prueba estadística de la sección 4.3 requiere que las dos muestras sean *independientes y aleatorias*. Ciertamente, el requisito de independencia es violado por la forma en la que se realizó el experimento.

El par de mediciones, en las llantas A y B, para un automóvil particular están definitivamente relacionadas. Una mirada a los datos muestra que las lecturas tienen más o menos la misma magnitud para un automóvil particular pero varían en forma marcada de un automóvil a otro. Esto, por supuesto, es exactamente lo que podría esperarse. El desgaste de llantas está determinado en su mayor parte por los hábitos de manejo, el balanceo de las ruedas y la superficie del pavimento. Como cada automóvil tiene un conductor diferente, es de esperarse una gran cantidad de variabilidad en los datos de un automóvil a otro.

Al diseñar el experimento de desgaste de llantas, el experimentador vio que las mediciones variarían en gran medida de un automóvil a otro. Si las llantas (cinco del tipo A y cinco del tipo B) se asignaran al azar a las 10 ruedas, resultando en muestras aleatorias *independientes*, esta variabilidad resultaría en un gran error estándar y haría difícil de detectar una diferencia en las medias. En cambio, el experimentador eligió "parear" las mediciones, comparando el desgaste para llantas tipo A y tipo B en cada uno de los cinco automóviles. Este diseño experimental, a veces llamado diseño de **diferencia pareada** o **pares acoplados**, nos permite eliminar la variabilidad de un auto a otro al observar sólo las cinco mediciones de diferencia mostradas en la tabla 4.4. Estas cinco diferencias forman una sola muestra aleatoria de tamaño $n = 5$.

TABLA 4.4 **Diferencias en desgaste de llantas, usando los datos de la tabla 4.3**

Automóvil	A	B	d = A − B
1	10.6	10.2	.4
2	9.8	9.4	.4
3	12.3	11.8	.5
4	9.7	9.1	.6
5	8.8	8.3	.5
			$\bar{d} = .48$

Observe que en la tabla 4.4 la media muestral de las diferencias, d = A – B, se calcula como

$$\bar{d} = \frac{\Sigma d_i}{n} = .48$$

y es exactamente la misma que la diferencia de las medias muestrales: $(\bar{x}_1 - \bar{x}_2) = (10.24 - 9.76) = .48$. No debe sorprender que esto se pueda demostrar como verdadero en general y también que la misma relación se cumpla para las medias poblacionales. De este modo, el promedio de las diferencias poblacionales es

$$\mu_d = (\mu_1 - \mu_2)$$

Debido a esto, se pueden usar las diferencias muestrales para probar una diferencia significativa en las dos medias poblacionales, $(\mu_1 - \mu_2) = \mu_d$. La prueba es una prueba t de una sola muestra de las mediciones de diferencia para probar la hipótesis nula

$$H_0 : \mu_d = 0 \quad [\text{o } H_0 : (\mu_1 - \mu_2) = 0]$$

contra la hipótesis alternativa

$$H_a : \mu_d \neq 0 \quad [\text{o } H_a : (\mu_1 - \mu_2) \neq 0]$$

Los procedimientos de prueba toman la misma forma que los procedimientos empleados en la sección 4.2 y se describen a continuación.

PRUEBA DE HIPÓTESIS DE DIFERENCIA PAREADA PARA $(\mu_1 - \mu_2) = \mu_d$: MUESTRAS DEPENDIENTES

1. Hipótesis nula: $H_0 : \mu_d = 0$
2. Hipótesis alternativa:

Prueba de una cola	Prueba de dos colas
$H_a: \mu_d > 0$	$H_a: \mu_d \neq 0$
(o $H_a: \mu_d < 0$)	

3. Estadístico de prueba: $t = \dfrac{\bar{d} - 0}{s_d/\sqrt{n}} = \dfrac{\bar{d}}{s_d/\sqrt{n}}$

donde n = Número de diferencias pareadas
$\bar{d}$ = Media de las diferencias muestrales
s_d = Desviación estándar de las diferencias muestrales

$$= \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{\Sigma d_i^2 - \dfrac{(\Sigma d_i)^2}{n}}{n-1}}$$

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	Prueba de dos colas
$t > t_\alpha$	$t > t_{\alpha/2}$ o $t < -t_{\alpha/2}$
(o $t < -t_\alpha$ cuando la hipótesis alternativa sea $H_a: \mu_d < 0$)	

o cuando valor $p < \alpha$

Los valores críticos de t, t_α y $t_{\alpha/2}$ están basados en $(n - 1)$ gl. Estos valores tabulados se pueden hallar usando la tabla 4 en el Apéndice.

INTERVALO DE CONFIANZA DE MUESTRA PEQUEÑA $(1 - \alpha)100\%$ PARA $(\mu_1 - \mu_2) = \mu_d$, CON BASE EN UN EXPERIMENTO DE DIFERENCIA PAREADA

$$\bar{d} \pm t_{\alpha/2}\left(\frac{s_d}{\sqrt{n}}\right)$$

Suposiciones: El experimento está diseñado como una prueba de diferencia pareada de modo que las n diferencias representan una muestra aleatoria tomada de una población normal.

EJEMPLO 4.8 ¿Los datos de la tabla 4.3 dan suficiente evidencia para indicar una diferencia en el desgaste medio para llantas tipo A y B? Pruebe usando $\alpha = .05$.

Solución Usted puede verificar, usando su calculadora, que el promedio y desviación estándar de las cinco mediciones de diferencia son

$$\bar{d} = .48 \quad \text{y} \quad s_d = \sqrt{\frac{1.18 - \dfrac{(2.4)^2}{5}}{4}} = .0837$$

Entonces

$$H_0: \mu_d = 0 \quad \text{y} \quad H_a: \mu_d \neq 0$$

y

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}} = \frac{.48}{.0837/\sqrt{5}} = 12.8$$

El valor crítico de t para una prueba estadística de dos colas, $\alpha = .05$ y 4 gl, es 2.776. Ciertamente, el valor observado de $t = 12.8$ es muy grande y significativo. En consecuencia, se puede concluir que hay una diferencia en el desgaste medio para llantas tipo A y B.

EJEMPLO 4.9

Encuentre un intervalo de confianza de 95% para $(\mu_1 - \mu_2) = \mu_d$ usando los datos de la tabla 4.3.

Solución Un intervalo de confianza de 95% para la diferencia entre la media de los niveles de desgaste es

$$\bar{d} \pm t_{\alpha/2}\left(\frac{s_d}{\sqrt{n}}\right)$$

$$.48 \pm 2.776\left(\frac{.0837}{\sqrt{5}}\right)$$

$$.48 \pm .10$$

o sea $.38 < (\mu_1 - \mu_2) < .58$. ¿Cómo se compara el ancho de este intervalo con el ancho de un intervalo que pudiera haberse construido *si* se hubiera diseñado el experimento en forma no pareada? Es probable que fuera de la misma magnitud que el intervalo calculado en la figura 4.11, donde los datos observados fueron *incorrectamente* analizados usando el análisis no pareado. Este intervalo, $-1.448 < (\mu_1 - \mu_2) < 2.408$, es mucho más ancho que el intervalo pareado, lo cual indica que el diseño de diferencia pareada aumentó la precisión de nuestra estimación y hemos obtenido valiosa información con el uso de este diseño.

TIP MI CONSEJO

¡Los intervalos de confianza siempre se interpretan en la misma forma! En el muestreo repetido, los intervalos construidos de esta manera encierran el valor verdadero del parámetro 100(1 α)% del tiempo

---

La *prueba de diferencia pareada* o *diseño de pares acoplados* empleados en el experimento de desgaste de llantas es un ejemplo sencillo de un diseño experimental denominado **diseño de bloque aleatorizado**. Cuando hay gran cantidad de variabilidad entre las unidades experimentales, aun antes de ponerse en práctica ningún procedimiento experimental, el efecto de esta variabilidad se puede reducir al mínimo al poner en **bloque**, es decir, comparar los diferentes procedimientos dentro de grupos de unidades experimentales relativamente similares llamadas **bloques**. De este modo, el "ruido" causado por la gran variabilidad no oculta las verdaderas diferencias entre los procedimientos. En el capítulo 6 analizaremos con más detalle los diseños aleatorizados de bloque.

Es importante que recuerde que el *pareo* o *poner en bloque* se presenta cuando se planea el experimento y no después de recolectar los datos. Un experimentador tiene la opción de usar pares de gemelos idénticos para comparar dos métodos de aprendizaje. Un médico puede registrar la presión sanguínea de un paciente antes y después de administrar un medicamento en particular. Una vez que usted haya empleado un diseño pareado para un experimento, ya no tiene la opción de usar el análisis no pareado de la sección 4.3. La suposición de independencia ha sido violada a propósito y la única opción es usar el análisis pareado descrito aquí.

TIP MI CONSEJO

Prueba de diferencia pareada: $df = n - 1$

Aun cuando hacer un pareado fue benéfico en el experimento de desgaste de llantas, éste puede no ser siempre el caso. En el análisis pareado, los grados de libertad para la prueba t se cortan a la mitad, es decir, de $(n + n - 2) = 2(n - 1)$ a $(n - 1)$. Esta reducción *aumenta* el valor crítico de t para rechazar H_0 y también aumenta el ancho del intervalo de confianza para la diferencia en las dos medias. Si el pareado no es eficiente, este aumento no es compensado por una disminución en la variabilidad y de hecho se puede perder y no ganar información por el

pareado. Esto por supuesto que no ocurrió en el experimento de las llantas, ya que la reducción grande en el error estándar compensó la pérdida en grados de libertad.

Excepto por la notación, el análisis de diferencia pareada es igual al análisis de una sola muestra presentado en la sección 4.2. No obstante, *MINITAB* y *MS Excel* dan un solo procedimiento (**Paired *t*** en *MINITAB* y ***t*–Test: Paired Two Sample for Means** en *MS Excel*) para analizar las diferencias. La salida de *MINITAB*, que se muestra en la figura 4.12, presenta el valor *p* para el análisis pareado, .000, que indica una diferencia altamente significativa en las medias. Encontrará instrucciones para generar salidas de *MINITAB* y de *MS Excel* en la sección 4.7.

FIGURA 4.12
Salida impresa *MINITAB* para análisis de diferencia pareada de datos del desgaste de llantas

Prueba T pareada y CI: llanta A, llanta B

```
Paired T for Tire A - Tire B
                    N       Mean         StDev    SE Mean
Tire A              5     10.240         1.316      0.589
Tire B              5      9.760         1.328      0.594
Difference          5     0.4800        0.0837     0.0374

95% CI for mean difference: (0.3761, 0.5839)
T-Test of mean difference = 0 (vs not = 0): T-Value = 12.83  P-Value = 0.000
```

4.5 INFERENCIAS ACERCA DE LA VARIANZA DE UNA POBLACIÓN

Ya hemos visto en las secciones precedentes que una estimación de la varianza poblacional σ^2 suele ser necesaria antes de hacer inferencias acerca de medias poblacionales, pero a veces la varianza poblacional σ^2 es el objetivo principal en una investigación experimental. ¡Ésta llega a ser *más* importante para el experimentador que la media poblacional! Considere estos ejemplos:

- Los instrumentos de mediciones científicas deben dar lecturas no sesgadas con un error de medición muy pequeño. Un altímetro de un avión que mida la altitud correcta en *promedio* es bastante inútil si las mediciones tienen un error de hasta 1 000 pies arriba o debajo de la altitud correcta.
- Las piezas maquinadas en un proceso de manufactura deben ser producidas con mínima variabilidad para reducir piezas de dimensiones incorrectas y, por tanto, defectuosas.
- Las pruebas de aptitud deben estar diseñadas de manera que las calificaciones *exhibirán* una cantidad razonable de variabilidad. Por ejemplo, un examen de 800 puntos no es muy discriminatorio si todos los estudiantes obtienen calificaciones entre 601 y 605.

En capítulos previos, hemos usado

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}$$

como estimador no sesgado de la varianza poblacional σ^2. Esto significa que, en muestreo repetido, el promedio de todas las estimaciones muestrales será igual al parámetro objetivo, σ^2. Pero, ¿qué tan cercano o lejano del objetivo es probable que esté su estimador s^2? Para contestar esta pregunta, usamos la distribución de muestreo de s^2, que describe su comportamiento en muestreo repetido.

Considere la distribución de s^2 basada en muestreo *aleatorio* repetido de una distribución *normal* con una media y varianza especificadas. Podemos demostrar teóricamente que la distribución empieza en $s^2 = 0$ (porque la varianza no puede ser negativa) con una media igual

a σ^2. Su forma es *no simétrica* y cambia con cada tamaño muestral diferente y cada valor diferente de σ^2. Hallar valores críticos para la distribución de muestreo de s^2 sería muy difícil y requeriría tablas separadas para cada varianza poblacional. Por fortuna, podemos simplificar el problema por *estandarización*, como hicimos con la distribución z.

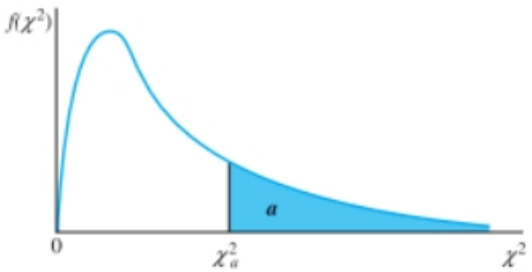
Definición La estadística estandarizada

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

recibe el nombre de **variable chi cuadrada** y tiene una distribución de muestreo llamada **distribución de probabilidad chi cuadrada**, con $n - 1$ grados de libertad.

La ecuación de la función de densidad para esta estadística es por demás complicada al verla, pero traza la curva que se muestra en la figura 4.13.

FIGURA 4.13
Distribución chi cuadrada



Ciertos valores críticos de la estadística chi cuadrada, que se usan para hacer inferencias acerca de la varianza poblacional, han sido tabulados por estadísticos y aparecen en la tabla 5 del Apéndice. Como la forma de la distribución varía con el tamaño muestral n o, con mayor precisión, los grados de libertad, $n - 1$, asociados con s^2, tabla 5, parcialmente reproducidos en la tabla 4.5, se construye exactamente en la misma forma que la tabla t, con los grados de libertad en la primera y última columnas. El símbolo χ^2_a indica que el valor χ^2 tabulado tiene un área a a su derecha (véase la figura 4.13).

TABLA 4.5 **Formato de la tabla chi cuadrada de la tabla 5 del Apéndice**

gl	$\chi^2_{.995}$	⋯	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	⋯	$\chi^2_{.005}$	gl
1	.0000393		.0039321	.0157908	2.70554	3.84146		7.87944	1
2	.0100251		.102587	.210720	4.60517	5.99147		10.5966	2
3	.0717212		.351846	.584375	6.25139	7.81473		12.8381	3
4	.206990		.710721	1.063623	7.77944	9.48773		14.8602	4
5	.411740		1.145476	1.610310	9.23635	11.0705		16.7496	5
6	.0675727		1.63539	2.204130	10.6446	12.5916		18.5476	6
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.
15	4.60094		7.26094	8.54675	22.3072	24.9958		32.8013	15
16	5.14224		7.96164	9.31223	23.5418	26.2962		34.2672	16
17	5.69724		8.67176	10.0852	24.7690	27.5871		35.7185	17
18	6.26481		9.39046	10.8649	25.9894	28.8693		37.1564	18
19	6.84398		10.1170	11.6509	27.2036	30.1435		38.5822	19
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.
.	.		.	.	.	.		.	.

TIP MI CONSEJO
Probando una varianza: $gl = n - 1$

MI APPLET EN LÍNEA
Probabilidades chi cuadrada

Se observa en la tabla 4.5 que, debido a que la distribución no es simétrica y empieza en 0, las áreas de cola superior e inferior deben ser tabuladas para la estadística chi cuadrada. Por ejemplo, el valor $\chi^2_{.95}$ es el valor que tiene 95% del área bajo la curva a su derecha y 5% del área a su izquierda. Este valor corta un área igual a .05 en la cola inferior de la distribución chi cuadrada.

EJEMPLO 4.10 Pruebe su capacidad para usar la tabla 5 del Apéndice al verificar las siguientes frases:

1. La probabilidad de que χ^2, basada en $n = 16$ mediciones ($gl = 15$), exceda de 24.9958 es .05.
2. Para una muestra de $n = 6$ mediciones, 95% del área bajo la distribución χ^2 está a la derecha de 1.145476.

Estos valores están sombreados en la tabla 4.5.

La prueba estadística de una hipótesis nula respecto a una varianza poblacional

$$H_0 : \sigma^2 = \sigma_0^2$$

usa el estadístico de prueba

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Observe que cuando H_0 es verdadera, s^2/σ_0^2 debe ser cercana a 1, de modo que χ^2 debe ser cercana a $(n-1)$, los grados de libertad. Si σ^2 es realmente mayor que el valor hipotético σ_0^2, el estadístico de prueba tenderá a ser mayor a $(n-1)$ y es probable que caiga hacia la cola superior de la distribución. Si $\sigma^2 < \sigma_0^2$, el estadístico de prueba tenderá a ser menor que $(n-1)$ y probablemente caerá hacia la cola inferior de la distribución chi cuadrada. Al igual que en otras situaciones de prueba, se puede usar una prueba estadística ya sea de una o de dos colas, dependiendo de la hipótesis alternativa. Esta prueba de hipótesis y el intervalo de confianza $(1-\alpha)100\%$ para σ^2 están basados ambos en la distribución chi cuadrada y se describen a continuación.

PRUEBA DE HIPÓTESIS RESPECTO A UNA VARIANZA POBLACIONAL

1. Hipótesis nula: $H_0 : \sigma^2 = \sigma_0^2$
2. Hipótesis alternativa:

Prueba de una cola	**Prueba de dos colas**
$H_a : \sigma^2 > \sigma_0^2$	$H_a : \sigma^2 \neq \sigma_0^2$
(o $H_a : \sigma^2 < \sigma_0^2$)	

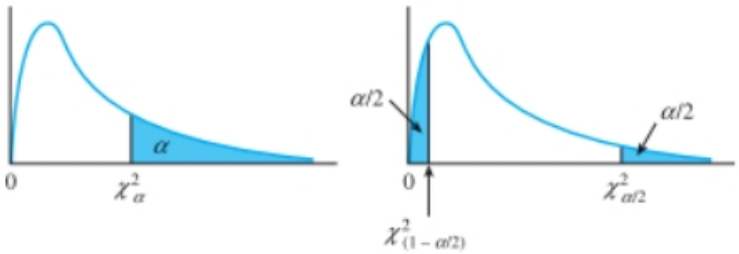
3. Estadístico de prueba: $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2}$

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	Prueba de dos colas
$\chi^2 > \chi^2_\alpha$ (o $\chi^2 < \chi^2_{(1-\alpha)}$ cuando la hipótesis alternativa sea $H_a: \sigma^2 < \sigma_0^2$), donde χ^2_α y $\chi^2_{(1-\alpha)}$ son, respectivamente, los valores de cola superior e inferior de χ^2 que pone α en las áreas de cola	$\chi^2 > \chi^2_{\alpha/2}$ o $\chi^2 < \chi^2_{(1-\alpha/2)}$, donde $\chi^2_{\alpha/2}$ y $\chi^2_{(1-\alpha/2)}$ son, respectivamente, los valores de cola superior e inferior de χ^2 que pone $\alpha/2$ en las áreas de cola

o cuando valor $p < \alpha$

Los valores críticos de χ^2 están basados en $(n - 1)$ *gl*. Estos valores tabulados se pueden hallar usando la tabla 5 del Apéndice.



INTERVALO DE CONFIANZA DE $(1 - \alpha)100\%$ PARA σ^2

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}}$$

donde $\chi^2_{(1-\alpha/2)}$ y $\chi^2_{\alpha/2}$ son los valores χ^2 superior e inferior, que localizan la mitad de α en cada cola de la distribución chi cuadrada.

Suposición: La muestra se selecciona al azar de una población normal.

EJEMPLO 4.11

Un fabricante de cemento declara que el concreto preparado con el producto de él tiene resistencia relativamente estable a la compresión y que la resistencia medida en kilogramos por centímetro cuadrado (kg/cm^2) está dentro de un rango de 40 kg/cm^2. Una muestra de $n = 10$ mediciones produjo una media y varianza igual a, respectivamente,

$$\bar{x} = 312 \quad \text{y} \quad s^2 = 195$$

¿Estos datos presentan suficiente evidencia para rechazar lo dicho por el fabricante?

Solución Usted ya sabe que el rango de un conjunto de mediciones debe ser aproximadamente cuatro desviaciones estándar. La afirmación del fabricante de que el rango de las mediciones de resistencia está dentro de 40 kg/cm^2 debe significar que la desviación estándar de las mediciones es casi 10 kg/cm^3 o menos. Para probar su declaración, las hipótesis apropiadas son

$$H_0: \sigma^2 = 10^2 = 100 \quad \text{contra} \quad H_a: \sigma^2 > 100$$

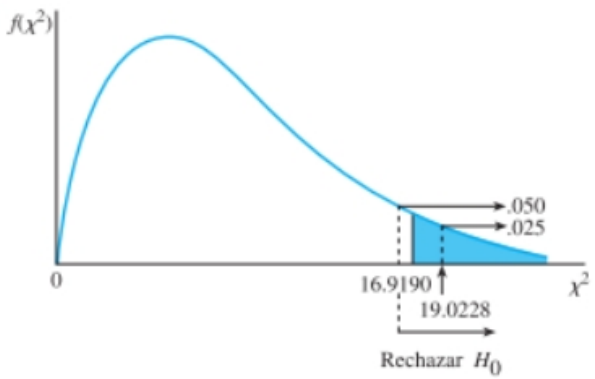
Si la varianza muestral es mucho mayor que el valor hipotético de 100, entonces el estadístico de prueba

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{1755}{100} = 17.55$$

será inusualmente grande, favoreciendo el rechazo de H_0 y la aceptación de H_a. Hay dos formas de usar el estadístico de prueba para tomar una decisión para esta prueba.

- **El método del valor crítico:** La prueba apropiada requiere una región de rechazo de una cola en la cola derecha de la distribución χ^2. El valor crítico para $\alpha = .05$ y $(n - 1) = 9$ *gl* es $\chi^2_{.05} = 16.9190$ de la tabla 5 en el Apéndice. La figura 4.14 muestra la región de rechazo; se puede rechazar H_0 si el estadístico de prueba excede de 16.9190. Como el valor observado del estadístico de prueba $\chi^2 = 17.55$, se puede concluir que la hipótesis nula es falsa y que el rango de mediciones de resistencia del concreto excede la declaración del fabricante.

FIGURA 4.14
Región de rechazo y valor *p* (sombreado) para el ejemplo 4.11



- **El método del valor *p***: El valor *p* para una prueba estadística es el mínimo valor de α para el cual H_0 puede ser rechazado. Se calcula, al igual que en otras pruebas de una cola, como el área en la cola de la distribución χ^2 a la derecha del valor observado, $\chi^2 = 17.55$. Aun cuando algunos paquetes de computadora permiten calcular esta área con toda exactitud, la tabla 5 del Apéndice permite sólo acotar el valor *p*. Como el valor 17.55 está entre $\chi^2_{.050} = 16.9190$ y $\chi^2_{.025} = 19.0228$, el valor *p* está entre .025 y .05. La mayoría de investigadores rechazarían H_0 y presentan estos resultados como significativos al nivel de 5%, o $P < .05$. De nueva cuenta, el investigador puede rechazar H_0 y concluir que el rango de mediciones excede de lo dicho por el fabricante.

EJEMPLO 4.12

Una experimentadora está convencida de que su instrumento de medición tenía una variabilidad medida por la desviación estándar $\sigma = 2$. Durante un experimento, ella registró las mediciones 4.1, 5.2 y 10.2. ¿Estos datos confirman o desaprueban lo dicho por ella? Pruebe la hipótesis apropiada y construya un intervalo de confianza de 90% para estimar el verdadero valor de la varianza de población.

Solución Como no hay nivel de significancia establecido de antemano, se debe elegir usar el método del valor *p* para probar estas hipótesis:

$$H_0 : \sigma^2 = 4 \quad \text{contra} \quad H_a : \sigma^2 \neq 4$$

Use su calculadora científica para verificar que la varianza muestral es $s^2 = 10.57$ y el estadístico de prueba es

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{2(10.57)}{4} = 5.285$$

Como ésta es una prueba de dos colas, la región de rechazo se divide en dos partes, la mitad en cada cola de la distribución χ^2. Si se aproxima el área a la derecha del estadístico de prueba observado, $\chi^2 = 5.285$, se tendrá sólo *la mitad* del valor p para la prueba. Como un valor igualmente improbable de χ^2 podría presentarse en la cola inferior de la distribución, con igual probabilidad, se debe *duplicar* el área superior para obtener el valor p. Con 2 *gl*, el valor observado, 5.29, cae entre $\chi^2_{.10}$ y $\chi^2_{.05}$ de modo que

$$.05 < \frac{1}{2}(\text{valor } p) < .10 \quad \text{o} \quad .10 < \text{valor } p < .20$$

Ya que el valor p es mayor que .10, los resultados no son estadísticamente significativos. Hay insuficiente evidencia para rechazar la hipótesis nula $H_0 : \sigma^2 = 4$.

El correspondiente intervalo de confianza de 90% es

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}}$$

Los valores de $\chi^2_{(1-\alpha/2)}$ y $\chi^2_{\alpha/2}$ son

$$\chi^2_{(1-\alpha/2)} = \chi^2_{.95} = .102587$$

$$\chi^2_{\alpha/2} = \chi^2_{.05} = 5.99147$$

Sustituyendo estos valores en la fórmula para la estimación de intervalo, se obtiene

$$\frac{2(10.57)}{5.99147} < \sigma^2 < \frac{2(10.57)}{.102587} \quad \text{o} \quad 3.53 < \sigma^2 < 206.07$$

Así, se puede estimar la varianza poblacional para que caiga en el intervalo 3.53 a 206.07. Este intervalo de confianza muy ancho indica la poca información sobre la varianza poblacional que se obtiene de una muestra de sólo tres mediciones. En consecuencia, no es de sorprenderse que haya insuficiente evidencia para rechazar la hipótesis nula $\sigma^2 = 4$. Para obtener más información sobre σ^2, el experimentador necesita aumentar el tamaño muestral.

Aunque *MS Excel* no tiene un comando único para implementar estos procedimientos, use la herramienta de función en *Excel* para encontrar usted mismo el estadístico de prueba, el valor p y los límites de confianza superior e inferior. Si usa *MINITAB*, el comando **Stat ▶ Basic Statistics ▶ 1 Variance** le permite introducir ya sea datos sin procesar o estadísticas de resumen para ejecutar la prueba chi cuadrada para una sola varianza, y calcular un intervalo de confianza. La parte pertinente de la salida impresa de *MINITAB 16* para el ejemplo 4.12 se muestra en la figura 4.15.

FIGURA 4.15
Salida de *MINITAB* para el ejemplo 4.12

Prueba y CI para una varianza: mediciones

```
Null hypothesis          Sigma-squared = 4
Alternative hypothesis   Sigma-squared not = 4

Statistics
Variable       N   StDev    Variance
Measurements   3    3.25        10.6

90% Confidence Intervals
CI for         CI for
Variable       Method          StDev          Variance
Measurements   Chi-Square   (1.88, 14.36)   (3.5, 206.1)
Test
Variable       Method         Statistic   DF   P-Value
Measurements   Chi-Square          5.28    2     0.142
```

4.6 COMPARACIÓN DE LAS VARIANZAS DE DOS POBLACIONES

Así como una sola varianza poblacional a veces es importante para un experimentador, quizás usted también necesite comparar dos varianzas poblacionales. Es probable que deba comparar la precisión de un aparato de medición con la de otro, la estabilidad de un proceso de manufactura con la de otro, o incluso la variabilidad en el procedimiento de calificaciones de un profesor universitario con el de otro.

Una forma de comparar dos varianzas poblacionales, σ_1^2 y σ_2^2, es usar la razón entre las varianzas muestrales, s_1^2/s_2^2. Si s_1^2/s_2^2 es casi igual a 1, se encuentra poca evidencia para indicar que σ_1^2 y σ_2^2 son diferentes. Por otra parte, un valor muy grande o muy pequeño para s_1^2/s_2^2 da evidencia de una diferencia en las varianzas poblacionales.

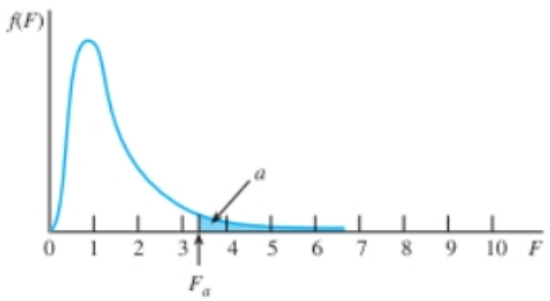
¿Qué tan grande debe ser s_1^2/s_2^2 para que haya evidencia suficiente para rechazar la siguiente hipótesis nula?

$$H_0 : \sigma_1^2 = \sigma_2^2$$

La respuesta a esta pregunta se halla al estudiar la distribución de s_1^2/s_2^2 en un muestreo repetido.

Cuando las muestras aleatorias independientes se sacan de dos poblaciones *normales* con *varianzas iguales*, es decir, $\sigma_1^2 = \sigma_2^2$; entonces s_1^2/s_2^2 tiene una distribución de probabilidad en un muestreo repetido que los estadísticos conocen como **distribución *F***. La ecuación de la función de densidad para este estadístico es bastante complicada para revisarla aquí, pero traza la curva que se muestra en la figura 4.16.

FIGURA 4.16
Una distribución F con $df_1 = 10$ y $df_2 = 10$



SUPOSICIONES PARA QUE s_1^2/s_2^2 TENGA UNA DISTRIBUCIÓN F

- Las muestras aleatorias e independientes se sacan de cada una de dos poblaciones normales.
- La variabilidad de las mediciones en las dos poblaciones es igual y puede ser medida por una varianza común, σ^2; esto es, $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

APPLET EN LÍNEA
Probabilidades F

No es importante que usted conozca la ecuación compleja de la función de densidad para F. Para nuestros fines, sólo se necesita usar los valores críticos bien tabulados de F dados en la tabla 6 en el Apéndice.

Al igual que la distribución χ^2, la forma de la distribución F es no simétrica y depende del número de grados de libertad asociados con s_1^2 y s_2^2, representados como $gl_1 = (n_1 - 1)$ y

TIP MI CONSEJO

Probar dos varianzas:
$df_1 = n_1 - 1$ y
$df_2 = n_2 - 1$

$gl_2 = (n_2 - 1)$, respectivamente. Esto complica la tabulación de valores críticos de la distribución F porque se requiere una tabla para cada combinación diferente de gl_1, gl_2 y a.

En la tabla 6 del Apéndice, los valores críticos de F para áreas de cola derecha correspondientes a $a = .100, .050, .025, .010$ y $.005$ están tabulados para varias combinaciones de gl_1 grados de libertad del numerador y gl_2 grados de libertad del denominador. Una parte de la tabla 6 se reproduce en la tabla 4.6. El número de grados de libertad gl_1 del numerador aparece en el margen superior y el número de grados de libertad del denominador gl_2 aparece a lo largo del margen lateral. Los valores de a se encuentran en la segunda columna. Para una combinación fija de gl_1 y gl_2, los valores críticos apropiados de F se encuentran en la línea indizada por el valor de a requerido.

EJEMPLO 4.13

Compruebe su capacidad para usar la tabla 6 del Apéndice al verificar los siguientes enunciados:

1. El valor de F con área .05 a su derecha para $gl_1 = 6$ y $gl_2 = 9$ es 3.37.
2. El valor de F con área .05 a su derecha para $gl_1 = 5$ y $gl_2 = 10$ es 3.33.
3. El valor de F con área .01 a su derecha para $gl_1 = 6$ y $gl_2 = 9$ es 5.80.

Estos valores están sombreados en la tabla 4.6.

TABLA 4.6 Formato para la tabla *F* de la tabla 6 del Apéndice

		gl_1					
gl_2	a	1	2	3	4	5	6
1	.100	39.86	49.50	53.59	55.83	57.24	58.20
	.050	161.4	199.5	215.7	224.6	230.2	234.0
	.025	647.8	799.5	864.2	899.6	921.8	937.1
	.010	4052	4999.5	5403	5625	5764	5859
	.005	16211	20000	21615	22500	23056	23437
2	.100	8.53	9.00	9.16	9.24	9.29	9.33
	.050	18.51	19.00	19.16	19.25	19.30	19.33
	.025	38.51	39.00	39.17	39.25	39.30	39.33
	.010	98.50	99.00	99.17	99.25	99.30	99.33
	.005	198.5	199.0	199.2	199.2	199.3	199.3
3	.100	5.54	5.46	5.39	5.34	5.31	5.28
	.050	10.13	9.55	9.28	9.12	9.01	8.94
	.025	17.44	16.04	15.44	15.10	14.88	14.73
	.010	34.12	30.82	29.46	28.71	28.24	27.91
	.005	55.55	49.80	47.47	46.19	45.39	44.84
.	.			.			.
.	.			.			.
.	.			.			.
9	.100	3.36	3.01	2.81	2.69	2.61	2.55
	.050	5.12	4.26	3.86	3.63	3.48	3.37
	.025	7.21	5.71	5.08	4.72	4.48	4.32
	.010	10.56	8.02	6.99	6.42	6.06	5.80
	.005	13.61	10.11	8.72	7.96	7.47	7.13
10	.100	3.29	2.92	2.73	2.61	2.52	2.46
	.050	4.96	4.10	3.71	3.48	3.33	3.22
	.025	6.94	5.46	4.83	4.47	4.24	4.07
	.010	10.04	7.56	6.55	5.99	5.64	5.39
	.005	12.83	9.43	8.08	7.34	6.87	6.54

4.7 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

Prueba y estimación de muestra pequeña: Microsoft Excel

Las pruebas de hipótesis para dos medias poblacionales basadas en la distribución *t* de Student y la prueba *F* para la razón de dos varianzas pueden encontrarse usando el comando de *Microsoft Excel* **Data ▶ Data Analysis.** Recuerde que necesita haber cargado los complementos de *Excel* llamados **Analysis ToolPak**. Encontrará tres opciones para las pruebas *t* de dos muestras y una prueba *F* en la lista de "Analysis Tools". Para elegir las pruebas *t* apropiadas, primero debe decidir si las muestras son independientes o pareadas; para la prueba de muestras independientes, debe decidir si es posible suponer que las varianzas poblacionales son iguales o no.

EJEMPLO 4.14 **(Prueba *t* de dos muestras suponiendo varianzas iguales)** Se registraron las puntuaciones de examen en la misma prueba de álgebra para nueve estudiantes seleccionados al azar de una clase instruida por el profesor A y ocho estudiantes seleccionados al azar de una clase enseñada por el profesor B. ¿Hay una diferencia en el promedio de las puntuaciones para estudiantes enseñados por estos dos profesores?

Profesor A	65	88	93	95	80	76	79	77	
Profesor B	91	85	70	82	92	68	86	87	75

Introduzca los datos en las columnas A y B de una hoja de cálculo de *Excel*.

1. Use **Data ▶ Data Analysis ▶ Descriptive Statistics** o **Statistical ▶ STDEV.S** (**STDEV** en *Excel 2007* y versiones anteriores) para hallar las desviaciones estándar para las dos muestras, $s_1 = 9.913$ y $s_2 = 8.800$. En vista que la razón de las dos varianzas es $s_1^2/s_2^2 = 1.27$ (menos de 3), se tiene seguridad al suponer que las varianzas poblacionales son iguales.
2. Seleccione **Data ▶ Data Analysis ▶ t–Test: Two–Sample Assuming Equal Variances** para generar el cuadro de diálogo que se muestra en la figura 10.19a. Resalte o teclee **Variable 1 Range** y **Variable 2 Range** (los datos en la primera y segunda columnas) en los primeros dos cuadros. En el cuadro marcado "Hypothesized Mean Difference" teclee **0** (en vista que se está probando H_0: $\mu_1 - \mu_2 = 0$) y seleccione "Labels" si es necesario.
3. En *Excel* el nivel de significancia predeterminado es $\alpha = .05$. Cambie este nivel de significancia si es necesario. Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK**. La salida aparecerá en la ubicación de la celda seleccionada y deberá ajustarse usando **Format ▶ AutoFit Column Width** en la pestaña **Home** en el grupo **Cells** mientras se encuentra resaltada todavía. Puede disminuir la precisión decimal si lo desea, usando en la pestaña **Home** en el grupo **Number**.
4. El valor observado del estadístico de prueba $t = -.0337$ se encuentra en la figura 10.19b en la fila etiquetada "*t Stat*" seguido por el valor *p* de una cola "*P(T<=t) one–tail*" y el valor crítico marcando la región de rechazo para una prueba de una cola con $\alpha = .05$. Las últimas dos filas de la salida proporcionan el valor *p* y el valor *t* crítico para una prueba de dos colas.
5. Para este ejemplo, el valor $p = .9736$ indica que no hay una diferencia significativa en el promedio de las puntuaciones para los estudiantes enseñados por los dos profesores.

FIGURA 4.17

(a)

(b)

D	E	F
t-Test: Two-Sample Assuming Equal Variances		
	A	*B*
Mean	81.625	81.778
Variance	98.268	77.444
Observations	8	9
Pooled Variance	87.162	
Hypothesized Mean Difference	0	
df	15	
t Stat	-0.0337	
P(T<=t) one-tail	0.4868	
t Critical one-tail	1.7531	
P(T<=t) two-tail	0.9736	
t Critical two-tail	2.1314	

6. En la sección 4.6 se presentó una prueba de hipótesis formal para la igualdad de las dos varianzas usando la prueba *F*. Para efectuar esta prueba usando *Excel*, seleccione **Data ▶ Data Analysis ▶ F–Test: Two–Sample for Variances**. Siga las instrucciones para la prueba *t* de varianzas iguales, pero reemplace el valor "Alpha" con 0.025, y generará la salida que se muestra en la figura 4.18.

FIGURA 4.18

F-Test Two-Sample for Variances		
	A	*B*
Mean	81.625	81.78
Variance	98.268	77.44
Observations	8	9
df	7	8
F	1.2689	
P(F<=f) one-tail	0.3701	
F Critical one-tail	4.5286	

Note que en la salida sólo se proporcionan el valor *p* de una cola y el valor crítico, razón por la cual se especifica que la cola única sea 0.025. Por tanto, para nuestra prueba de dos colas, $\alpha = 0.05$ y el valor $p = .7402$. No hay una diferencia significativa en las dos varianzas.

EJEMPLO 4.15

(Prueba *t* de dos muestras suponiendo varianzas desiguales)

1. Consulte el ejemplo 4.14. Si la razón de las dos varianzas muestrales ha sido tan grande que no podría suponer varianzas iguales (se usa "mayor a 3" como regla general), debería seleccionar **Data ▶ Data Analysis ▶ t–Test: Two–Sample Assuming Unequal Variances**.
2. Siga las instrucciones para la prueba *t* de varianzas iguales, y generará una salida similar. Si se usa esta prueba en los datos del ejemplo 4.14, resulta la siguiente salida (figura 4.19).

FIGURA 4.19

H	I	J
t-Test: Two-Sample Assuming Unequal Variances		
	A	*B*
Mean	81.625	81.778
Variance	98.268	77.444
Observations	8	9
Hypothesized Mean Difference	0	
df	14	
t Stat	-0.0334	
P(T<=t) one-tail	0.4869	
t Critical one-tail	1.7613	
P(T<=t) two-tail	0.9738	
t Critical two-tail	2.1448	

3. Verá ligeras diferencias en el valor observado del estadístico de prueba, los grados de libertad y los valores *p* para la prueba, pero las conclusiones no cambian.
4. NOTA: Cuando calcula los grados de libertad para la *Aproximación de Satterthwaite*, la **Data Analysis Tool** en *Excel* redondea al entero más cercano. Una función alternativa de *Excel* para calcular el valor *p* para esta prueba (▶ **Statistical** ▶ **T.TEST**) usa el valor exacto de *gl* dados por la fórmula de Satterthwaite. Debido a estos enfoques diferentes para determinar los grados de libertad, los resultados de **T.TEST** y la herramienta *t*–Test diferirán ligeramente en el caso de varianzas desiguales, y *también* diferirán ligeramente de la salida de *MINITAB*.

EJEMPLO 4.16 **(Prueba *t* pareada)** Consulte los datos del desgaste de las llantas de la tabla 4.3.

1. Para ejecutar una prueba de diferencia pareada para estas muestras dependientes, introduzca los datos en las primeras dos columnas de una hoja de cálculo de Excel y seleccione **Data** ▶ **Data Analysis** ▶ **t–Test: Paired Two Sample for Means**.
2. Siga las instrucciones para la prueba *t* de varianzas iguales, y generará una salida similar. Para los datos de la tabla 4.3, obtendrá la salida que se muestra en la figura 4.20. Una vez más, puede disminuir la precisión decimal si lo desea, usando en la pestaña **Home** en el grupo **Number**.
3. Usando el valor observado del estadístico de prueba ($t = 12.83$) con valor *p* de dos colas = .0002, hay una evidencia fuerte que indica una diferencia en las dos medias poblacionales.

FIGURA 4.20

D	E	F
t-Test: Paired Two Sample for Means		
	Tire A	*Tire B*
Mean	10.24	9.76
Variance	1.733	1.763
Observations	5	5
Pearson Correlation	0.998	
Hypothesized Mean Difference	0	
df	4	
t Stat	12.8285	
P(T<=t) one-tail	0.0001	
t Critical one-tail	2.1318	
P(T<=t) two-tail	0.0002	
t Critical two-tail	2.7764	

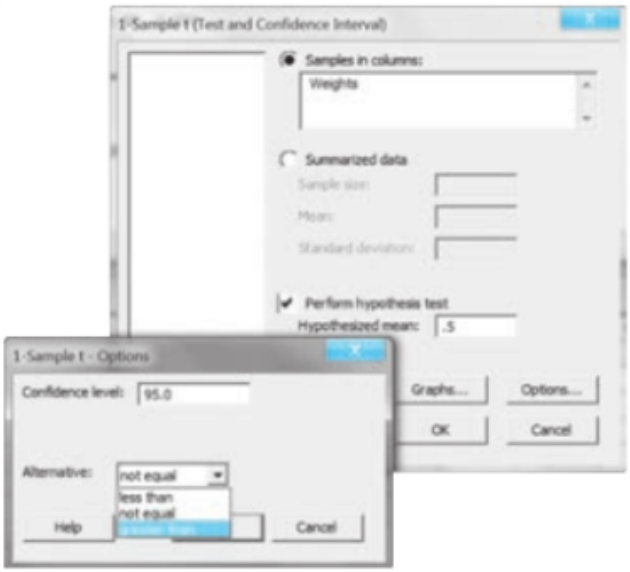
Prueba y estimación de muestra pequeña: *MINITAB*

Las pruebas de hipótesis para dos medias poblacionales basadas en la distribución *t* de Student y la prueba *F* para la razón de dos varianzas pueden encontrarse usando el comando de *MINITAB* **Stat** ▶ **Basic Statistics**. Encontrará opciones para **1–Sample t**, **2–Sample t**, **Paired t** y **2 Variances**, las cuales ejecutarán las pruebas y procedimientos de estimación de las secciones 4.2, 4.3, 4.4 y 4.6. Para elegir las pruebas *t* de dos muestras apropiadas, primero debe decidir si las muestras son independientes o pareadas; para la prueba de muestras independientes, debe decidir si es posible suponer que las varianzas poblacionales son iguales o no.

EJEMPLO 4.17 **(Prueba *t* de una muestra)** Consulte el ejemplo 4.3, en el cual se comparó el promedio del peso de diamantes usando un proceso nuevo con el promedio de un peso de .5 quilates.

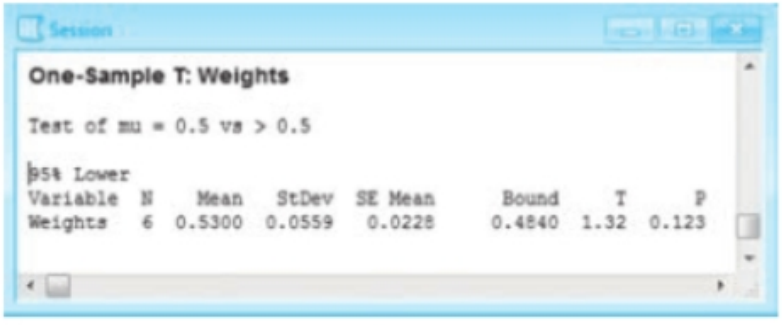
1. Introduzca los seis pesos registrados: .46, .61, .52, .48, .57, .54, en la columna C1 y nómbrela "Pesos". Use **Stat ▶ Basic Statistics ▶ 1–Sample t** para generar los cuadros de diálogo que se muestran en la figura 4.21.

FIGURA 4.21



2. Para probar H_0: $\mu = .5$ contra H_a: $\mu > .5$, use la lista de la izquierda para seleccionar "Pesos" para el cuadro marcado "Samples in Columns". Seleccione el cuadro marcado "Perform hypothesis test". Luego, coloque el cursor en el cuadro marcado "Hypothesized mean:" e introduzca **.5** como el valor de prueba. Por último, use **Options** y el menú desplegable marcado "Alternative" para seleccionar "greater than". Puede cambiar el coeficiente de confianza por omisión de **.95** si lo desea. Haga clic en **OK** dos veces para obtener la salida que se muestra en la figura 4.22.

FIGURA 4.22



3. Note que *MINITAB* produce un intervalo de confianza de una o dos colas para la media poblacional única, consistente con la hipótesis alternativa que haya elegido.

EJEMPLO 4.18 **(Prueba *t* de dos muestras)** Se registraron las puntuaciones de examen en la misma prueba de álgebra para nueve estudiantes seleccionados al azar de un salón de clase enseñado por el profesor A y ocho estudiantes seleccionados al azar de un salón de clase enseñado por el profesor B. ¿Hay una diferencia en el promedio de las puntuaciones para estudiantes enseñados por estos dos profesores?

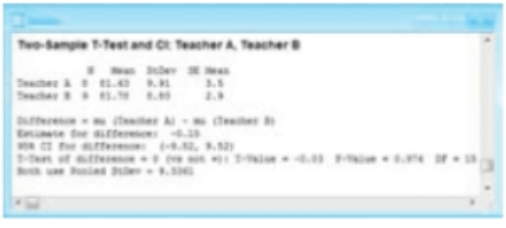
Profesor A	65	88	93	95	80	76	79	77	
Profesor B	91	85	70	82	92	68	86	87	75

1. Los datos pueden introducirse en la hoja de cálculo en una de tres formas:
   - Introduzca las mediciones de ambas muestras en una sola columna e introduzca las letras (A o B) en una segunda columna para identificar la muestra de la cual proviene la medición.
   - Introduzca las muestras en dos columnas separadas.
   - Si no tiene los datos sin procesar, pero en su lugar tiene estadísticas de resumen, *MINITAB 16* le permitirá usar estos valores al seleccionar "Summarized data" e introducir los valores apropiados en los cuadros.
2. Utilice el segundo método e introduzca los datos en dos columnas de la hoja de cálculo. Use **Stat ▶ Basic Statistics ▶ Display Descriptive Statistics** para hallar las desviaciones estándar para las dos muestras, $s_1 = 9.91$ y $s_2 = 8.80$. En vista que la razón de las dos varianzas es $s_1^2/s_2^2 = 1.27$ (menor que 3), se tiene seguridad al suponer que las varianzas poblacionales son iguales.
3. Seleccione **Stat ▶ Basic Statistics ▶ 2–Sample t** para generar el cuadro de diálogo que se muestra en la figura 4.23a. Seleccione "samples in different columns", seleccionando las columnas apropiadas de la lista de la izquierda. Elija el cuadro "Assume equal variances" y seleccione la alternativa apropiada en el cuadro **Options**. La salida de dos muestras aparece cuando haga clic dos veces en **OK** y automáticamente contiene un intervalo de confianza del 95% de una o dos colas al igual que el estadístico de prueba y el valor *p* (puede cambiar el coeficiente de confianza si lo desea). La salida se muestra en la figura 4.23b.

FIGURA 4.23

(a)

(b)



4. El valor observado del estadístico de prueba $t = -.03$ es etiquetado "*T–Value*" seguido por el "*P–Value*" de dos colas. Para este ejemplo, el valor $p = .974$ indica que no hay una diferencia significativa en el promedio de las puntuaciones para los estudiantes enseñados por los dos profesores.
5. En la sección 4.6 se presentó una prueba de hipótesis formal para la igualdad de dos varianzas usando la prueba *F*. Para ejecutar esta prueba usando *MINITAB*, seleccione **Stat ▶ Basic Statistics ▶ 2 Variances**. En la lista desplegable, seleccione "Samples in two columns" e introduzca las columnas apropiadas de la lista de la izquierda. La porción pertinente de la salida se muestra en la figura 4.24. Para nuestra prueba de dos colas con $\alpha = 0.05$, el estadístico de prueba es $F = 1.27$ y el valor $p = .740$. No hay una diferencia significativa en las dos varianzas.

FIGURA 4.24

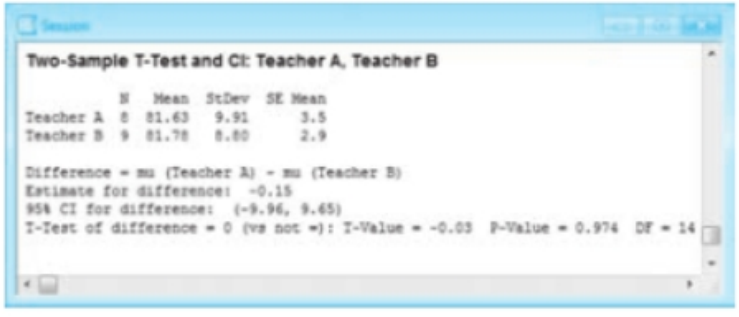
```
Test and CI for Two Variances: Teacher A, Teacher B

Null hypothesis         Sigma(Teacher A) / Sigma(Teacher B) = 1
Alternative hypothesis  Sigma(Teacher A) / Sigma(Teacher B) not = 1
Significance level      Alpha = 0.05

Test
Method                          DF1  DF2  Statistic  P-Value
F Test (normal)                   7    8       1.27    0.740
```

EJEMPLO 4.19

(Prueba *t* de dos muestras suponiendo varianzas desiguales)

1. Consulte el ejemplo 4.18. Si la razón de las dos varianzas muestrales hubiera sido tan grande que no pudiera suponer varianzas iguales (se usa "mayor que 3" como regla general), debería seleccionar **Stat ▶ Basic Statistics ▶ 2–Sample t**, pero NO seleccione el cuadro marcado "Assume Equal Variances". Si usamos esta prueba en los datos del ejemplo 4.14, resulta la siguiente salida (figura 4.25).

FIGURA 4.25



```
Two-Sample T-Test and CI: Teacher A, Teacher B

           N   Mean  StDev  SE Mean
Teacher A  8  81.63   9.91      3.5
Teacher B  9  81.78   8.80      2.9

Difference = mu (Teacher A) - mu (Teacher B)
Estimate for difference:  -0.15
95% CI for difference:  (-9.96, 9.65)
T-Test of difference = 0 (vs not =): T-Value = -0.03  P-Value = 0.974  DF = 14
```

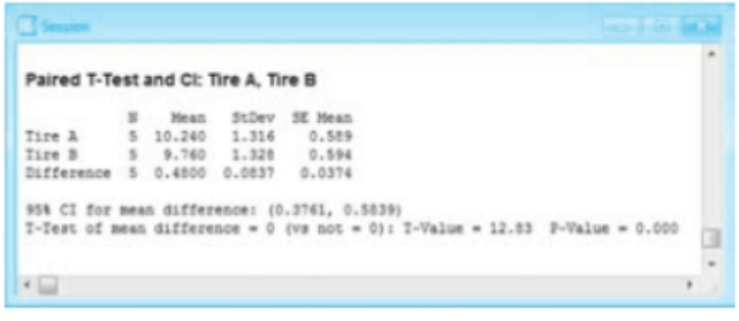
2. Verá ligeras diferencias en los grados de libertad y no hay lista "Pooled StDev", pero las conclusiones no cambian.
3. NOTA: Cuando *MINITAB* calcula los grados de libertad para la *Aproximación de Satterthwaite*, *MINITAB* usa la parte entera del valor calculado, el cual es diferente de los procedimientos usados en *MS Excel*. Debido a estos enfoques diferentes para determinar los grados de libertad, los resultados de las salidas de *MS Excel* y *MINITAB* diferirán ligeramente.

EJEMPLO 4.20

(Prueba de diferencia pareada) Consulte los datos del desgaste de las llantas de la tabla 4.3.

1. Para realizar una prueba de diferencia pareada para estas muestras dependientes, introduzca los datos en las primeras dos columnas de una hoja de cálculo de *MINITAB* y seleccione **Stat ▶ Basic Statistics ▶ Paired t**.
2. Siga las instrucciones para la prueba *t* de muestras independientes, y generará una salida similar. Para los datos en la tabla 4.3, se obtiene la salida en la figura 4.26.
3. Usando el valor observado del estadístico de prueba ($t = 12.83$) con un valor p de dos colas $= .000$, hay una evidencia fuerte que indica una diferencia en las dos medias poblacionales.

FIGURA 4.26



```
Paired T-Test and CI: Tire A, Tire B

            N    Mean   StDev  SE Mean
Tire A      5  10.240   1.316    0.589
Tire B      5   9.760   1.328    0.594
Difference  5  0.4800  0.0837   0.0374

95% CI for mean difference: (0.3761, 0.5839)
T-Test of mean difference = 0 (vs not = 0): T-Value = 12.83  P-Value = 0.000
```

EJERCICIOS

TÉCNICAS BÁSICAS

4.1 Encuentre los siguientes valores t en la tabla 4 del Apéndice:

a. $t_{.05}$ para 5 *gl*.
b. $t_{.025}$ para 8 *gl*.
c. $t_{.10}$ para 18 *gl*.
d. $t_{.025}$ para 30 *gl*.

4.2 Encuentre el (los) valor(es) crítico(s) de t que especifiquen la región de rechazo en estas situaciones:

a. Una prueba de dos colas con $\alpha = .01$ y 12 *gl*.
b. Una prueba de cola derecha con $\alpha = .05$ y 16 *gl*.
c. Una prueba de dos colas con $\alpha = .05$ y 25 *gl*.
d. Una prueba de cola izquierda con $\alpha = .01$ y 7 *gl*.

4.3 Las siguientes $n = 10$ observaciones son una muestra de una población normal:

7.4 7.1 6.5 7.5 7.6 6.3 6.9 7.7 6.5 7.0

a. Encuentre la media y desviación estándar de estos datos.
b. Encuentre un límite superior de confianza de una cola, de 99%, para la media poblacional μ.
c. Pruebe H_0: $\mu = 7.5$ contra H_a: $\mu < 7.5$. Use $\alpha = .01$.
d. ¿Los resultados de la parte **b** apoyan sus conclusiones de la parte **c**?

4.4 Calcule s^2, el estimador agrupado de σ^2, en estos casos:

a. $n_1 = 10, n_2 = 4, s_1^2 = 3.4, s_2^2 = 4.9$
b. $n_1 = 12, n_2 = 21, s_1^2 = 18, s_2^2 = 23$

4.5 Muestras aleatorias independientes de $n_1 = 16$ y $n_2 = 13$ observaciones fueron seleccionadas de dos poblaciones normales con iguales varianzas:

	Población	
	1	2
Tamaño muestral	16	13
Media muestral	34.6	32.2
Varianza muestral	4.8	5.9

a. Suponga que usted desea detectar una diferencia entre las medias poblacionales. Exprese las hipótesis nula y alternativa para la prueba.
b. Encuentre la región de rechazo para la prueba de la parte **a** para $\alpha = .01$.
c. Encuentre el valor del estadístico de prueba.
d. Encuentre el valor p aproximado para la prueba.
e. Realice la prueba y exprese sus conclusiones.

4.6 Se realizó un experimento de diferencia pareada usando $n = 10$ pares de observaciones.

a. Pruebe la hipótesis nula H_0: $(\mu_1 - \mu_2) = 0$ contra H_a: $(\mu_1 - \mu_2) \neq 0$ para $\alpha = .05$, $\bar{d} = .3$ y $s_d^2 = .16$. Dé el valor p aproximado para la prueba.
b. Encuentre un intervalo de confianza de 95% para $(\mu_1 - \mu_2)$.
c. ¿Cuántos pares de observaciones se necesitan si se desea estimar $(\mu_1 - \mu_2)$ correcta a no más de .1 con probabilidad igual a .95?

4.7 Una muestra aleatoria de tamaño $n = 7$ de una población normal produjo estas mediciones: 1.4, 3.6, 1.7, 2.0, 3.3, 2.8, 2.9.

a. Calcule la varianza muestral, s^2.
b. Construya un intervalo de confianza de 95% para la varianza poblacional, σ^2.
c. Pruebe H_0: $\sigma^2 = .8$ contra H_a: $\sigma^2 \neq .8$ usando $\alpha = .05$. Exprese sus conclusiones.
d. ¿Cuál es el valor p aproximado para la prueba de la parte **c**?

4.8 Una muestra aleatoria de $n = 12$ observaciones de una población normal produjo $\bar{x} = 47.1$ y $s^2 = 4.7$. Pruebe la hipótesis H_0: $\mu = 48$ contra H_0: $\mu \neq 48$ en un nivel de significancia del 5%.

APLICACIONES

DATOS EX1006

4.9 Atunes ¿Hay diferencia en los precios del atún, dependiendo del método de empaque? *Consumer Reports* da el promedio del precio estimado para una lata de 6 onzas o bolsa de 7.06 onzas de atún, con base en los precios pagados a nivel nacional en supermercados.[1] Estos precios están registrados para varias marcas diferentes de atún.

Atún claro en agua		Atún blanco en aceite	Atún blanco en agua	Atún claro en aceite	
.99	.53	1.27	1.49	2.56	.62
1.92	1.41	1.22	1.29	1.92	.66
1.23	1.12	1.19	1.27	1.30	62
.85	.63	1.22	1.35	1.79	.65
.65	.67		1.29	1.23	.60
.69	.60		1.00		.67
.60	.66		1.27		
			1.28		

Fuente: Estudio práctico "Pricing of Tuna" Copyright 2001 por Consumers Union of U.S. Inc., Yonkers, NY 10703-1057, organización sin fines de lucro. Reimpreso con permiso de la edición de junio de 2001 de *Consumer Report*® sólo con fines educativos. No se permite uso comercial ni reproducción. www.ConsumerReports.org®.

Suponga que las marcas de atún incluidas en este estudio representan una muestra aleatoria de todas las marcas de atún disponibles en Estados Unidos.

a. Encuentre un intervalo de confianza de 95% para el promedio del precio para el atún claro en agua. Interprete este intervalo. Esto es, ¿a qué se refiere ese "95%"?
b. Encuentre un intervalo de confianza de 95% para el precio promedio del atún blanco en aceite. ¿Cómo se compara el ancho de este intervalo con el ancho del intervalo de la parte **a**? ¿Puede explicar por qué?
c. Encuentre intervalos de confianza de 95% para las otras dos muestras (atún blanco en agua y atún claro en aceite). Grafique las cuatro medias de tratamiento y sus errores estándar en una gráfica bidimensional semejante a la figura 3.5. ¿Qué clase de comparaciones generales pueden hacerse acerca de los cuatro tratamientos? (Analizaremos el procedimiento para comparar más de dos medias poblacionales en el capítulo 6.)

DATOS EX1030

4.10 Pasitas Se contó el número de pasitas en cada una de 14 minicajas (tamaño de 1/2 onza), tanto de una marca genérica de pasitas como de la marca Sunmaid®.

Marca genérica				Sunmaid			
25	26	25	28	25	29	24	24
26	28	28	27	28	24	28	22
26	27	24	25	25	28	30	27
26	26			28	24		

a. Aun cuando las cantidades tienen una distribución normal, ¿estos datos tienen distribuciones aproximadamente normales? (SUGERENCIA: Use un histograma o gráfica de tallo y hoja.)

b. ¿Está usted dispuesto a suponer que las varianzas poblacionales subyacentes son iguales? ¿Por qué?

c. Use el método del valor p para determinar si hay una diferencia significativa en la media de los números de pasitas por minicaja. ¿Cuáles son las implicaciones de su conclusión?

DATOS EX1041

4.11 Canasta de mercado de Estados Unidos Un anuncio para una popular cadena de supermercados dice que había tenido precios consistentemente más bajos que uno de sus competidores. Como parte de una encuesta realizada por una compañía independiente de revisión de precios, el promedio total semanal, con base en los precios de alrededor de 95 artículos, se da para esta cadena y para su competidor, registrado durante cuatro semanas consecutivas en un mes en particular.

Semana	Anunciante ($)	Competidor ($)
1	254.26	256.03
2	240.62	255.65
3	231.90	255.12
4	234.13	261.18

a. ¿Hay una diferencia significativa en el promedio de los precios para estas dos cadenas diferentes de supermercados?

b. ¿Cuál es el valor p aproximado para la prueba realizada en la parte **a**?

c. Construya un intervalo de confianza de 99% para la diferencia en el promedio de los precios para las dos cadenas de supermercados. Interprete este intervalo.

DATOS EX1048

4.12 Música en el trabajo Antes de contratar la instalación de música estéreo transmitida a cada uno de los espacios de oficinas, un ejecutivo hizo que su gerente seleccionara al azar siete oficinas para instalarles el sistema. El promedio del tiempo (en minutos) empleado fuera de estas oficinas, por salidas de los oficinistas involucrados, se registró antes y después de instalar el sistema de música con los siguientes resultados:

Número de oficina	1	2	3	4	5	6	7
Sin música	8	9	5	6	5	10	7
Con música	5	6	7	5	6	7	8

¿Sugeriría usted que el ejecutivo proceda con la instalación? Realice una prueba apropiada de hipótesis. Encuentre el valor p aproximado e interprete sus resultados.

4.13 Precisión de instrumentos Un instrumento de precisión está garantizado para leer con exactitud con variación de no más de 2 unidades. Una muestra de cuatro lecturas de instrumentos en el mismo objeto dio las mediciones 353, 351, 351 y 355.

a. Pruebe la hipótesis nula de que $\sigma = .7$ contra la alternativa de $\sigma < .7$. Use $\alpha = .05$.

b. Encuentre un intervalo de confianza del 90% para la varianza poblacional.

4.14 Pino rojo El crecimiento del tallo principal, medido para una muestra de 17 árboles de pino rojo de cuatro años de edad, produjo una media y desviación estándar igual a 11.3 y 3.4 pulgadas, respectivamente. Encuentre un intervalo de confianza de 90% para el crecimiento medio de una población de árboles de pino rojo de cuatro años de edad sujetos a condiciones ambientales similares.

DATOS EX1071

4.15 Alcohol y tiempos de reacción Para probar el efecto del alcohol al aumentar el tiempo de reacción para responder a un estímulo dado, se midieron los tiempos de reacción de siete personas. Después de consumir 3 onzas de alcohol al 40%, el tiempo de reacción para cada una de las siete personas se midió otra vez. ¿Los datos siguientes indican que la media de tiempo de reacción después de consumir alcohol fue mayor que la media de tiempo de reacción antes de consumir alcohol? Use $\alpha = .05$.

Persona	1	2	3	4	5	6	7
Antes	4	5	5	4	3	6	2
Después	7	8	3	5	4	5	5

4.16 ¿En casa o en preescolar? Cuatro pares de gemelos idénticos (pares A, B, C y D) se seleccionaron al azar de una base de datos computarizada de gemelos idénticos. Un niño se seleccionó al azar de cada par para formar un "grupo experimental". Estos cuatro niños fueron enviados a preescolar. Los otros cuatro niños se mantuvieron en casa como grupo de control. Al final del año escolar, se obtuvieron las siguientes calificaciones de CI (cociente de inteligencia):

Par	Grupo experimental	Grupo de control
A	110	111
B	125	120
C	139	128
D	142	135

¿Esta evidencia justifica la conclusión de que la falta de experiencia preescolar tiene un efecto desalentador en calificaciones del CI? Use el método del valor p.

DATOS EX1094

4.17 Tiempos de reacción Un psicólogo desea verificar que cierto medicamento aumenta el tiempo de reacción a cierto estímulo. Los siguientes tiempos de reacción (en décimas de segundo) se registraron antes y después de la inyección del medicamento para cada una de las cuatro personas:

	Tiempo de reacción	
Persona	Antes	Después
1	7	13
2	2	3
3	12	18
4	12	13

Pruebe al nivel de significancia de 5% para determinar si el medicamento aumentó significativamente el tiempo de reacción.

4.18 El sueño y el estudiante universitario ¿Cuánto duerme usted en una noche escolar típica? A un grupo de 10 estudiantes universitarios se le pidió informar el número de horas que dormían en la noche previa, con los siguientes resultados:

7, 6, 7.25, 7, 8.5, 5, 8, 7, 6.75, 6

a. Encuentre un intervalo de confianza de 99% para el promedio del número de horas que duermen los estudiantes universitarios.

b. ¿Qué suposiciones se requieren para que este intervalo de confianza sea válido?

Fuente

1. "Pricing of Tuna," *Consumer Reports*, junio de 2001.

5

Prueba de hipótesis

Scott Olson/Getty Images

OBJETIVOS GENERALES

En este capítulo, el concepto de una prueba estadística de hipótesis se presenta de manera formal. Las distribuciones muestrales de estadísticas que se estudiaron en los capítulos 2 y 3 se usan para construir pruebas de muestra grande respecto a los valores de parámetros poblacionales de interés para el experimentador.

ÍNDICE DEL CAPÍTULO

- Prueba de muestras grandes alrededor de una media poblacional μ
- Prueba de muestras grandes alrededor de $(\mu_1 - \mu_2)$
- Prueba de una hipótesis alrededor de $(p_1 - p_2)$
- Prueba de una hipótesis alrededor de una proporción poblacional p
- Una prueba estadística de hipótesis

NECESITO SABER...

Regiones de rechazo, valores *p* y conclusiones
Cómo calcular β

5.1 DEFINICIÓN, PRUEBA DE HIPÓTESIS DE UNO Y DOS EXTREMOS

En situaciones prácticas, una inferencia estadística comprende la estimación de un parámetro poblacional o tomar decisiones acerca del valor del parámetro. Por ejemplo, si una compañía farmacéutica fermenta un tanque de antibiótico se pueden usar muestras del tanque para *estimar* la potencia media μ para todo el antibiótico del tanque. En contraste, suponga que la compañía no se interesa en la potencia media exacta del antibiótico, sino sólo en satisfacer los estándares de potencia mínimos del gobierno. Entonces la compañía puede utilizar muestras del tanque para decidir entre estas dos posibilidades:

- La potencia media μ no excede la potencia mínima permisible.
- La potencia media μ excede la potencia mínima permisible.

El problema de la compañía farmacéutica ilustra una prueba estadística de hipótesis.

El razonamiento empleado en una **prueba estadística de hipótesis** es similar al proceso en un tribunal. Al procesar a una persona por robo, el tribunal debe decidir entre inocencia y culpabilidad. Cuando el juicio inicia se supone que la persona acusada es *inocente*. La fiscalía recaba y presenta toda evidencia disponible en un intento para contradecir la hipótesis de inocencia y por tanto obtener una condena.

Si hay evidencia suficiente contra la inocencia, el tribunal rechazará la hipótesis de inocencia y declarará *culpable* al demandado. Si la fiscalía no presenta suficiente evidencia para demostrar que el demandado es culpable, el tribunal le hallará *no culpable*. Observe que esto no demuestra que el demandado es inocente, sino sólo que no hubo evidencia suficiente para concluir que el demandado era culpable.

Empleamos este tipo de razonamiento para explicar los conceptos básicos de prueba de hipótesis. Estos conceptos se utilizan para probar los cuatro parámetros poblacionales expuestos en el capítulo 3: una sola media o proporción poblacional (μ o p) y la diferencia entre dos medias o proporciones poblacionales ($\mu_1 - \mu_2$ o $p_1 - p_2$). Cuando los tamaños muestrales son grandes, los estimadores puntuales para cada uno de estos cuatro parámetros tienen distribuciones muestrales normales, de modo que las cuatro pruebas estadísticas de muestras grandes siguen el mismo modelo general.

5.2 TIPOS DE HIPÓTESIS: NULA Y ALTERNATIVA

Una prueba estadística de hipótesis está formada por cinco partes:

1. La hipótesis nula, denotada por H_0
2. La hipótesis alternativa, denotada por H_a
3. El estadístico de prueba y su valor p
4. La región de rechazo
5. La conclusión

Cuando se especifican estos cinco elementos se define una prueba particular; el cambio de una o más de las partes crea una nueva prueba. Veamos con más detalle cada parte de la prueba estadística de hipótesis.

Definición Las dos hipótesis en competencia son la **hipótesis alternativa** H_a, generalmente la que el investigador desea apoyar, y **la hipótesis nula** H_0, una contradicción de la hipótesis alternativa.

1–2

Como veremos, es más fácil presentar apoyo para la hipótesis alternativa al demostrar que la hipótesis nula es falsa. En consecuencia, el investigador estadístico siempre empieza por suponer que la hipótesis nula H_0 es verdadera y utiliza entonces los datos muestrales para decidir si la evidencia está a favor de H_a más que de H_0 y saca una de dos **conclusiones**:

- Rechaza H_0 y concluye que H_a es verdadera.
- Acepta (no rechaza) H_0 como verdadera.

EJEMPLO 5.1

Se desea demostrar que el promedio de salario por hora de carpinteros en el estado de California es diferente de $19, que es el promedio nacional. Ésta es la hipótesis alternativa, escrita como

2 $H_a : \mu \neq 19$

La hipótesis nula es

1 $H_0 : \mu = 19$

A usted le gustaría rechazar la hipótesis nula, con lo que concluiría que la media en California no es igual a $19.

EJEMPLO 5.2

Un proceso de fresado produce un promedio de 3% de piezas defectuosas. Usted está interesado en demostrar que un simple ajuste en una máquina reducirá p, la proporción de piezas defectuosas producidas en el proceso de fresado. Entonces, la hipótesis alternativa es

2 $H_a : p < .03$

y la hipótesis nula es

1 $H_0 : p = .03$

Si puede rechazar H_0, se concluye que el proceso ajustado produce menos de 3% de piezas defectuosas.

5.3 NIVEL DE SIGNIFICANCIA, ERRORES TIPO I Y TIPO II

Si el estadístico de prueba cae en la región de rechazo, entonces se rechaza la hipótesis nula. Si el estadístico de prueba cae en la región de aceptación, entonces la hipótesis nula se acepta o la prueba se juzga como no concluyente. Aclararemos los diferentes tipos de conclusiones que son apropiados cuando consideremos varios ejemplos prácticos de prueba de hipótesis.

Por último, ¿cómo decide sobre los **valores críticos** que separan las regiones de aceptación y rechazo? Es decir, ¿cómo decide cuánta evidencia estadística necesita antes de rechazar H_0? Esto depende de la cantidad de confianza que usted, el investigador, desea añadir a las conclusiones de prueba y el **nivel de significancia** α, el riesgo que está dispuesto a correr si se toma una decisión incorrecta.

Definición Un **error tipo I** para una prueba estadística es el error de rechazar la hipótesis nula cuando es verdadera. El **nivel de significancia** para una prueba estadística de hipótesis es

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar falsamente } H_0) = P(\text{rechazar } H_0 \text{ cuando es verdadera})$$

Este valor α representa el *máximo riesgo tolerable* de rechazar incorrectamente H_0. Una vez fijo este nivel de significancia, la región de rechazo se establece para permitir que el investigador rechace H_0 con un grado fijo de confianza en la decisión.

Como hay dos opciones en una prueba estadística, también hay dos tipos de errores que se pueden cometer. En la sala de juzgado el acusado podría ser considerado no culpable cuando en realidad es culpable, o viceversa; lo mismo es cierto en una prueba estadística. De hecho, la hipótesis nula puede ser verdadera o falsa, sin considerar la decisión que tome el experimentador. Estas dos posibilidades, junto con las dos decisiones que puede tomar el investigador, se muestran en la tabla 5.1.

TABLA 5.1 **Tabla de decisión**

	Hipótesis nula	
Decisión	Verdadera	Falsa
Rechazar H_0	Error tipo I	Decisión correcta
Aceptar H_0	Decisión correcta	Error tipo II

Además del error tipo I con probabilidad α definida antes en esta sección, es posible cometer un segundo error, llamado **error tipo II**, el cual tiene probabilidad β.

Definición Un **error tipo I** para una prueba estadística ocurre si se rechaza la hipótesis nula cuando es verdadera. La probabilidad de cometer un error tipo I se denota por el símbolo α.

Un **error tipo II** para una prueba estadística ocurre si se acepta la hipótesis nula cuando es falsa y alguna hipótesis alternativa es verdadera. La probabilidad de cometer un error tipo II se denota por el símbolo β.

Observe que la probabilidad de un error tipo I es exactamente igual que el **nivel de significancia α** y, por tanto, es controlada por el investigador. Cuando H_0 es rechazada, se tiene una medida precisa de la confiabilidad de la inferencia; la probabilidad de una decisión incorrecta es α, pero la probabilidad β de un error tipo II no siempre es controlada por el experimentador. De hecho, cuando H_0 es falsa y H_a es verdadera, puede que no sea posible especificar un valor exacto para μ, sino sólo un rango de valores. Esto hace difícil, si no imposible, calcular β. Sin una medida de confiabilidad, no es inteligente concluir que H_0 sea verdadera. En lugar de arriesgarse a una decisión incorrecta, debería retener el juicio, concluyendo que *no hay evidencia suficiente para rechazar* H_0. En lugar de *aceptar* H_0, debe *"no rechazar" o "fallar en rechazar"* H_0.

TIP MI CONSEJO

$\alpha = P$(rechazar H_0 cuando H_0 es verdadera)

$\beta = P$(aceptar H_0 cuando H_0 es falsa)

Recuerde que *"aceptar" una hipótesis particular significa decidir en su favor*. Cualquiera que sea el resultado de una prueba, nunca se está *seguro* que la hipótesis que se "acepte" es verdadera. *Siempre hay un riesgo de estar equivocado* (*medido por α o β*). En consecuencia, nunca "acepte" H_0 si β es desconocida o su valor es inaceptable para usted. Cuando ocurra esta situación, debería retener el juicio y recolectar más información.

5.4 PRUEBA DE HIPÓTESIS PARA MEDIAS Y PROPORCIONES

En numerosas situaciones, la pregunta estadística que debe ser contestada involucra una comparación de dos medias poblacionales. Por ejemplo, el Servicio Postal de Estados Unidos está interesado en reducir su gasto de gasolina al cambiar sus camiones de motor de gasolina por

camiones eléctricos. Para determinar si se obtienen ahorros importantes en costos de operación al cambiar a camiones eléctricos, debe efectuarse un estudio piloto usando, por ejemplo, cien camiones convencionales del correo con motor de gasolina y cien camiones eléctricos operados en condiciones similares.

El estadístico que resume la información muestral respecto a la diferencia en medias poblacionales $(\mu_1 - \mu_2)$ es la diferencia en medias muestrales $(\bar{x}_1 - \bar{x}_2)$. Por tanto, al probar si la diferencia en medias muestrales indica que la diferencia verdadera en medias poblacionales difiere de un valor especificado, $(\mu_1 - \mu_2) = D_0$, se puede usar el error estándar de $(\bar{x}_1 - \bar{x}_2)$,

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \text{estimada por} \quad \text{SE} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

en la forma de un estadístico z para medir a cuántas desviaciones estándar se encuentra la diferencia $(\bar{x}_1 - \bar{x}_2)$ desde la diferencia hipotética D_0. A continuación se describe el procedimiento formal de prueba.

PRUEBA ESTADÍSTICA DE MUESTRAS GRANDES PARA $(\mu_1 - \mu_2)$

1. Hipótesis nula: H_0: $(\mu_1 - \mu_2) = D_0$, donde D_0 es alguna diferencia especificada que desea probar. Para muchas pruebas, usted hará hipótesis de que no hay diferencia entre μ_1 y μ_2; esto es, $D_0 = 0$.
2. Hipótesis alternativa:

Prueba de una cola	**Prueba de dos colas**
$H_a : (\mu_1 - \mu_2) > D_0$ [o $H_a : (\mu_1 - \mu_2) < D_0$]	$H_a : (\mu_1 - \mu_2) \neq D_0$

3. Estadístico de prueba: $z \approx \dfrac{(\bar{x}_1 - \bar{x}_2) - D_0}{\text{SE}} = \dfrac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$
4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	**Prueba de dos colas**
$z > z_\alpha$ [o $z < -z_\alpha$ cuando la hipótesis alternativa es H_a: $(\mu_1 - \mu_2) < D_0$] o cuando el valor $p < \alpha$	$z > z_{\alpha/2}$ o bien $z < -z_{\alpha/2}$

Suposiciones: Las muestras son seleccionadas al azar y de manera independiente de las dos poblaciones $n_1 \geq 30$ y $n_2 \geq 30$.

EJEMPLO 5.3 Para determinar si ser propietario de un automóvil afecta el rendimiento académico de un estudiante, se tomaron dos muestras aleatorias de 100 estudiantes de sexo masculino. El promedio de calificaciones para los $n_1 = 100$ no propietarios de un automóvil tuvieron un promedio y varianza igual a $\bar{x}_1 = 2.70$ y $s_1^2 = .36$, en tanto que $\bar{x}_2 = 2.54$ y $s_2^2 = .40$ para los $n_2 = 100$

propietarios de un automóvil. ¿Los datos presentan suficiente evidencia para indicar una diferencia en la media en el rendimiento entre propietarios y no propietarios de un automóvil? Pruebe usando $\alpha = .05$.

Solución Para detectar una diferencia, si existe, entre la media de los rendimientos académicos para no propietarios μ_1 y los propietarios de un automóvil μ_2, probaremos la hipótesis nula de que no hay diferencia entre las medias contra la hipótesis alternativa de que $(\mu_1 - \mu_2) \neq 0$; esto es,

$$H_0: (\mu_1 - \mu_2) = D_0 = 0 \qquad \text{contra} \qquad H_a: (\mu_1 - \mu_2) \neq 0$$

Sustituyendo en la fórmula para el estadístico de prueba, obtenemos

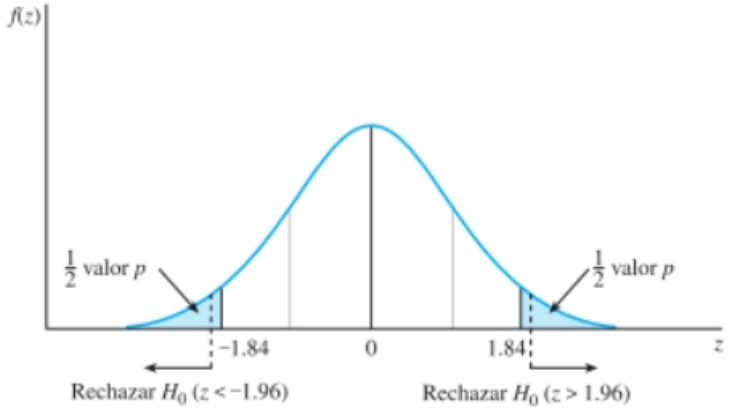
$$z \approx \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2.70 - 2.54}{\sqrt{\frac{.36}{100} + \frac{.40}{100}}} = 1.84$$

- **El método del valor crítico:** Usando una prueba de dos colas con nivel de significancia $\alpha = .05$, se pone $\alpha/2 = .025$ en cada cola de la distribución z y se rechaza H_0 si $z > 1.96$ o $z < -1.96$. Como $z = 1.84$ no excede de 1.96 y no es menor que -1.96, H_0 no puede ser rechazada (véase la figura 5.1). Esto es, no hay evidencia suficiente para declarar una diferencia en el promedio de los rendimientos académicos para los dos grupos. Recuerde que no debe estar dispuesto a *aceptar* H_0, es decir, declarar que las dos medias son iguales, hasta que β sea evaluada para algunos valores significativos de $(\mu_1 - \mu_2)$.

TIP MI CONSEJO

$|$Estadístico de prueba$|$ $>$ $|$valor crítico$|$ $\Leftrightarrow$ rechazar H_0

FIGURA 5.1
Región de rechazo y valor p para el ejemplo 5.3



- **El método del valor p:** Calcule el valor p, la probabilidad de que z es mayor que $z = 1.84$ más la probabilidad de que z sea menor que $z = -1.84$, como se muestra en la figura 5.1:

$$\text{Valor } p = P(z > 1.84) + P(z < -1.84) = (1 - .9671) + .0329 = .0658$$

El valor p se encuentra entre .10 y .05, de modo que se puede rechazar H_0 al nivel .10 pero no al nivel de significancia .05. Como el valor p de .0658 excede del nivel de significancia especificado $\alpha = .05$, H_0 no puede ser rechazada. De nuevo, usted no debería estar dispuesto a *aceptar* H_0 hasta que β sea evaluada para algunos valores significativos de $(\mu_1 - \mu_2)$.

Prueba de hipótesis e intervalos de confianza

Si usa el método del valor crítico o del valor p para probar hipótesis acerca de $(\mu_1 - \mu_2)$, siempre llegará a la misma conclusión porque el valor calculado del estadístico de prueba y el valor crítico están relacionados *exactamente* en la misma forma que están relacionados el valor p

y el nivel de significancia α. Hay que recordar que los intervalos de confianza construidos en el capítulo 3 podrían también usarse para contestar preguntas acerca de la diferencia entre dos medias poblacionales. De hecho, para una prueba de dos colas, el intervalo de confianza $(1 - \alpha)100\%$ para el parámetro de interés se usa para probar su valor, igual que como hicimos de manera informal en el capítulo 3.

El valor de α indicado por el coeficiente de confianza en el intervalo de confianza es equivalente al nivel de significancia α en el estadístico de prueba. Además, con el método del intervalo de confianza, se gana un margen de posibles valores para el parámetro de interés, cualquiera que sea el resultado de la prueba de hipótesis.

- Si el intervalo de confianza que se construye *contiene* el valor del parámetro especificado por H_0, entonces ese valor es uno de los posibles valores del parámetro y H_0 no debe ser rechazada.
- Si el valor hipotético *se encuentra fuera* de los límites de confianza, la hipótesis nula es rechazada al nivel de significancia α.

EJEMPLO 5.4 Construya un intervalo de confianza de 95% para la diferencia en el promedio de rendimiento académico entre propietarios y no propietarios de automóvil. Usando el intervalo de confianza, ¿se puede concluir que hay una diferencia en las medias poblacionales para los dos grupos de estudiantes?

Solución Consulte la sección 3.3. Para la diferencia en dos medias poblacionales, el intervalo de confianza se aproxima como

$$(\bar{x}_1 - \bar{x}_2) \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$(2.70 - 2.54) \pm 1.96\sqrt{\frac{.36}{100} + \frac{.40}{100}}$$

$$.16 \pm .17$$

o bien, $-.01 < (\mu_1 - \mu_2) < .33$. Este intervalo da un margen de posibles valores para la diferencia en las medias poblacionales.

Como la diferencia hipotética, $(\mu_1 - \mu_2) = 0$, está contenida en el intervalo de confianza, no se debería rechazar H_0. Vea los signos de los posibles valores del intervalo de confianza. No puede decir por el intervalo si la diferencia en las medias es negativa (–), positiva (+) o cero (0), donde este último indicaría que las dos medias son iguales. En consecuencia, realmente no se puede llegar a una conclusión en términos de la pregunta planteada. No hay suficiente evidencia para indicar que haya diferencia en el promedio de los rendimientos académicos para propietarios y no propietarios de automóvil. La conclusión es la misma que en el ejemplo 5.3.

Prueba para una proporción binomial

Cuando una muestra aleatoria de n intentos idénticos se saca de una población binomial, la proporción muestral $\hat{p}$ tiene una distribución aproximadamente normal cuando n es grande, con media p y error estándar

$$\text{SE} = \sqrt{\frac{pq}{n}}$$

Cuando se prueba una hipótesis acerca de p, la proporción en la población que posee cierto atributo, la prueba sigue la misma forma general que las pruebas de muestra grande de la sección 5.4. Para probar una hipótesis de la forma

$$H_0: p = p_0$$

contra una alternativa de una o de dos colas

$$H_a: p > p_0 \quad \text{o} \quad H_a: p < p_0 \quad \text{o} \quad H_a: p \neq p_0$$

el estadístico de prueba se construye usando $\hat{p}$, el mejor estimador de la verdadera proporción poblacional p. La proporción muestral $\hat{p}$ es estandarizada, usando la media y el error estándar hipotéticos, para formar una estadística de prueba z, que tiene una distribución normal estándar si H_0 es verdadera. Esta prueba de muestra grande se resume a continuación.

PRUEBA ESTADÍSTICA DE MUESTRAS GRANDES PARA p

1. Hipótesis nula: $H_0: p = p_0$
2. Hipótesis alternativa:

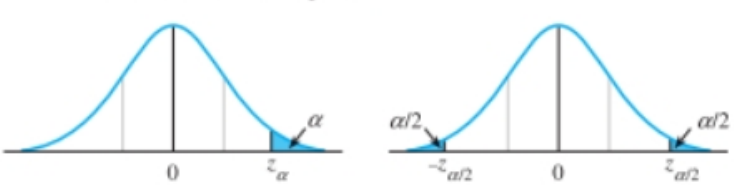
Prueba de una cola	**Prueba de dos colas**
$H_a: p > p_0$	$H_a: p \neq p_0$
(o, $H_a: p < p_0$)	

3. Estadístico de prueba: $z = \dfrac{\hat{p} - p_0}{\text{SE}} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$ con $\hat{p} = \dfrac{x}{n}$

   donde x es el número de éxitos en n intentos binomiales.[†]
4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	**Prueba de dos colas**
$z > z_\alpha$ (o $z < -z_\alpha$ cuando la hipótesis alternativa sea $H_a: p < p_0$)	$z > z_{\alpha/2}$ o bien $z < -z_{\alpha/2}$

   o cuando el valor $p < \alpha$



Suposición: El muestreo satisface las suposiciones de un experimento binomial y n es lo suficientemente grande para que la distribución muestral de $\hat{p}$ pueda ser aproximada por una distribución normal ($np_0 > 5$ y $nq_0 > 5$).

EJEMPLO 5.5 A cualquier edad, alrededor de 20% de los adultos estadounidenses participa en actividades de acondicionamiento físico al menos dos veces a la semana. No obstante, estas actividades cambian conforme las personas se hacen mayores, y ocasionalmente, los participantes dejan de serlo al envejecer. En una encuesta local de $n = 100$ adultos de más de

[†] Un estimador equivalente se obtiene al multiplicar el numerador y el denominador por z entre n, por lo que

$$z = \frac{x - np_0}{\sqrt{np_0 q_0}}$$

40 años, un total de 15 personas indicaron que participaron en estas actividades al menos dos veces a la semana. ¿Estos datos indican que el porcentaje de participación para adultos de más de 40 años de edad es considerablemente menor a la cifra de 20%? Calcule el valor p y úselo para sacar las conclusiones apropiadas.

Solución Suponiendo que el procedimiento de muestreo satisfaga los requisitos de un experimento binomial, se puede contestar la pregunta planteada usando una prueba de hipótesis de una cola:

$$H_0: p = .2 \quad \text{contra} \quad H_a: p < .2$$

TIP MI CONSEJO

Valor $p \leq \alpha \Leftrightarrow$ rechazar H_0

Valor $p > \alpha \Leftrightarrow$ no rechazar H_0

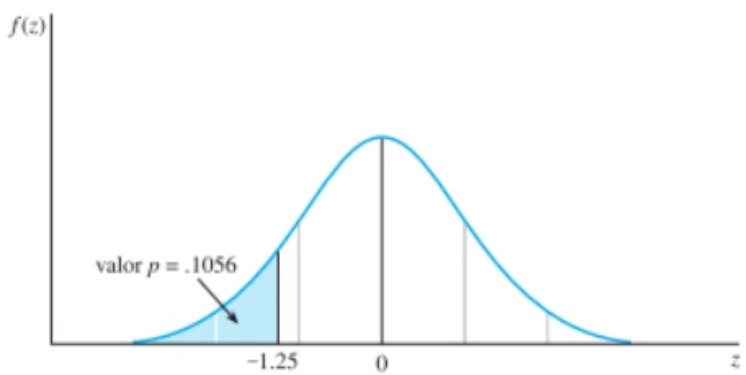
Empiece por suponer que H_0 es verdadera, es decir, el verdadero valor de p es $p_0 = .2$. Entonces $\hat{p} = x/n$ tendrá una distribución normal aproximada con media p_0 y error estándar $\sqrt{p_0q_0/n}$. (NOTA: Esto es diferente del procedimiento de estimación en el que el error estándar desconocido es estimado por $\sqrt{\hat{p}\hat{q}/n}$.) El valor observado de $\hat{p}$ es 15/100 = .15 y el estadístico de prueba es

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0q_0}{n}}} = \frac{.15 - .20}{\sqrt{\dfrac{(.20)(.80)}{100}}} = -1.25$$

El valor p asociado con esta prueba se encuentra como el área bajo la curva normal estándar a la izquierda de $z = -1.25$ como se observa en la figura 5.2. Por tanto,

$$\text{Valor } p = P(z < -1.25) = .1056$$

FIGURA 5.2
Valor p para el ejemplo 5.5



Si usted usa las guías para evaluar valores p, entonces .1056 es mayor que .10, y no rechazaría H_0. No hay evidencia suficiente para concluir que el porcentaje de adultos de más de 40 años que participan en actividades de acondicionamiento físico dos veces a la semana es menor a 20%.

Significancia estadística e importancia práctica

Es importante entender la diferencia entre resultados que sean "significativos" y resultados que son prácticamente "importantes". En lenguaje estadístico, la palabra *significativo* no necesariamente quiere decir "importante", sino sólo que los resultados podrían no haber ocurrido por casualidad. Suponga que en el ejemplo 5.5, la investigadora había empleado $n = 400$ adultos en su experimento y había observado la misma proporción muestral. El estadístico de prueba es ahora

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0q_0}{n}}} = \frac{.15 - .20}{\sqrt{\dfrac{(.20)(.80)}{400}}} = -2.50$$

con

$$\text{Valor } p = P(z < -2.50) = .0062$$

Ahora los resultados son *altamente significativos*: H_0 es rechazada y hay suficiente evidencia para indicar que el porcentaje de adultos de más de 40 años que participan en actividades de acondicionamiento físico es menor que 20%. No obstante, ¿esta baja en actividad es realmente *importante*? Supongamos que los médicos estarían preocupados sólo por una baja en actividad física de más de 10%. Si hubiera habido una caída mayor que 10% en actividad física, esto implicaría que el verdadero valor de p era menor que .10. ¿Cuál es el máximo valor posible de p? Usando un límite de confianza superior de 95% de una cola, tendremos

$$\hat{p} + 1.645\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$.15 + 1.645\sqrt{\frac{(.15)(.85)}{400}}$$

$$.15 + .029$$

o $p < .179$. La actividad física para adultos de 40 años o más ha bajado de 20% pero no se puede decir que haya caído por debajo de 10%. Entonces, los resultados, aun cuando son *estadísticamente significativos*, no son *prácticamente importantes*.

En este libro, usted aprenderá a determinar si los resultados son estadísticamente significativos. Sin embargo, cuando use estos procedimientos en una situación práctica, también debe asegurarse que los resultados sean prácticamente importantes.

Prueba para la diferencia entre dos proporciones binomiales

Cuando se seleccionan muestras aleatorias e independientes de dos poblaciones *binomiales*, el punto central del experimento puede ser la diferencia $(p_1 - p_2)$ en las proporciones de individuos u objetos que poseen una característica especificada en las dos poblaciones. En esta situación, usted puede usar la diferencia en las proporciones muestrales $(\hat{p}_1 - \hat{p}_2)$ junto con su error estándar,

$$\text{SE} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

en la forma de un estadístico z para probar una diferencia significativa en las dos proporciones poblacionales. La hipótesis nula a probarse suele ser de la forma

$$H_0: p_1 = p_2 \qquad \text{o} \qquad H_0: (p_1 - p_2) = 0$$

contra una hipótesis alternativa ya sea de una cola o de dos colas. La prueba formal de hipótesis se resume en el siguiente recuadro. Al estimar el error estándar para el estadístico z, se debe usar el hecho de que cuando H_0 es verdadera, las dos proporciones poblacionales son iguales a algún valor común, por ejemplo p. Para obtener la mejor estimación de este valor común, los datos muestrales son "agrupados" y la estimación de p es

$$\hat{p} = \frac{\text{Número total de éxitos}}{\text{Número total de pruebas}} = \frac{x_1 + x_2}{n_1 + n_2}$$

Recuerde que, para que la diferencia de las proporciones muestrales tengan una distribución aproximadamente normal, los tamaños muestrales deben ser grandes y las proporciones no deben estar demasiado cerca de 0 o 1.

TIP MI CONSEJO

Recuerde: Cada intento resulta en uno de dos resultados (S o F)

PRUEBA ESTADÍSTICA DE MUESTRAS GRANDES PARA $(p_1 - p_2)$

1. Hipótesis nula: H_0: $(p_1 - p_2) = 0$ o equivalentemente H_0: $p_1 = p_2$
2. Hipótesis alternativa:

Prueba de una cola	**Prueba de dos colas**
H_a: $(p_1 - p_2) > 0$ [o H_a: $(p_1 - p_2) < 0$]	H_a: $(p_1 - p_2) \neq 0$

3. Estadístico de prueba: $$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\text{SE}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{pq}{n_1} + \frac{pq}{n_2}}}$$

   donde $\hat{p}_1 = x_1/n_1$ y $\hat{p}_2 = x_2/n_2$. Como el valor común de $p_1 = p_2 = p$ (utilizado en el error estándar) se desconoce, se estima con

   $$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

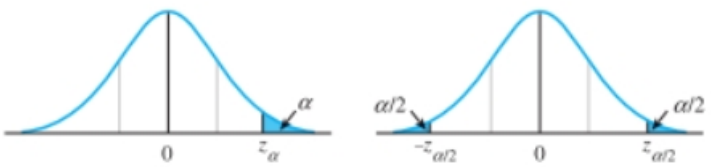
   y el estadístico de prueba es

   $$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}} \quad \text{o} \quad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	**Prueba de dos colas**
$z > z_\alpha$ [o $z < -z_\alpha$ cuando la hipótesis alternativa es H_a: $(p_1 - p_2) < 0$]	$z > z_{\alpha/2}$ o bien $z < -z_{\alpha/2}$

   o cuando el valor $p < \alpha$



Suposiciones: Las muestras se seleccionan al azar y de manera independiente provenientes de dos poblaciones binomiales, n_1 y n_2 son lo suficientemente grandes para que la distribución muestral de $(\hat{p}_1 - \hat{p}_2)$ pueda ser aproximada por una distribución normal. Esto es, $n_1\hat{p}_1$, $n_1\hat{q}_1$, $n_2\hat{p}_2$ y $n_2\hat{q}_2$ deberían ser todas mayores que 5.

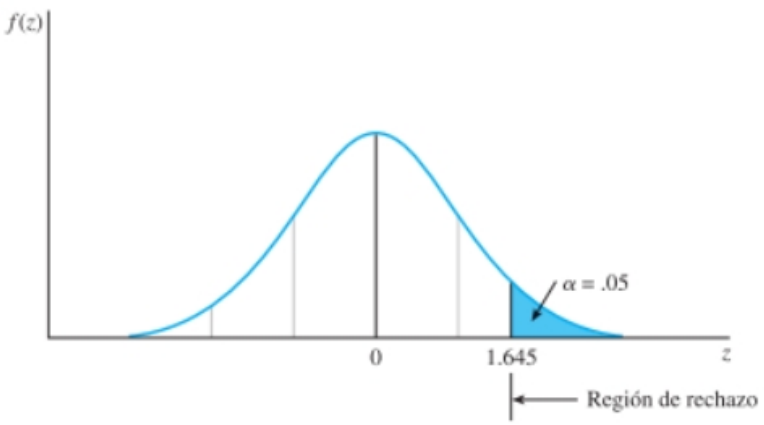
EJEMPLO 5.6 Los registros de un hospital indican que 52 hombres de una muestra de 1 000 contra 23 mujeres de una muestra de 1 000 fueron ingresados por enfermedad del corazón. ¿Estos datos presentan suficiente evidencia para indicar un porcentaje más alto de enfermedades del corazón entre hombres ingresados al hospital? Use $\alpha = .05$.

Solución Suponga que el número de pacientes ingresados por enfermedad del corazón tiene una distribución aproximada binomial de probabilidad para hombres y mujeres con parámetros p_1 y p_2, respectivamente. Entonces, como usted desea determinar si $p_1 > p_2$, probará la hipótesis nula $p_1 = p_2$, es decir, H_0: $(p_1 - p_2) = 0$, contra la hipótesis alternativa H_a: $p_1 > p_2$ o bien, lo que es equivalente, H_a: $(p_1 - p_2) > 0$. Para efectuar esta prueba, use el estadístico de prueba z y aproxime el error estándar usando la estimación agrupada de p. Como H_a implica una prueba de una cola, puede rechazar H_0 sólo para valores grandes de z. Por tanto, para $\alpha = .05$, puede rechazar H_0 si $z > 1.645$ (véase la figura 5.3).

La estimación agrupada de p requerida para el error estándar es

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{52 + 23}{1000 + 1000} = .0375$$

FIGURA 5.3
Ubicación de la región de rechazo para el ejemplo 5.6



y el estadístico de prueba es

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{.052 - .023}{\sqrt{(.0375)(.9625)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} = 3.41$$

Como el valor calculado de z cae en la región de rechazo, puede rechazar la hipótesis de que $p_1 = p_2$. Los datos presentan suficiente evidencia para indicar que el porcentaje de hombres que ingresan al hospital por enfermedad del corazón es más alto que el de mujeres. (NOTA: Esto no implica que la *incidencia* de enfermedad del corazón sea más alta en hombres. ¡Quizá menos mujeres ingresen al hospital cuando están afectadas por esa enfermedad!)

¿Cuánto *más alta* es la proporción de hombres que de mujeres que ingresan al hospital con enfermedad del corazón? Un límite inferior de una cola de 95% de confianza ayudará a hallar el mínimo valor probable para la diferencia.

$$(\hat{p}_1 - \hat{p}_2) - 1.645\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(.052 - .023) - 1.645\sqrt{\frac{.052(.948)}{1000} + \frac{.023(.977)}{1000}}$$

$$.029 - .014$$

o $(p_1 - p_2) > .015$. La proporción de hombres es aproximadamente 1.5% más alta que de mujeres. ¿Esto es de *importancia práctica*? Ésta es una pregunta que debe contestar el investigador.

En algunas situaciones necesitará probar para una diferencia D_0 (que no sea 0) entre dos proporciones binomiales. Si éste es el caso, el estadístico de prueba se modifica para probar H_0: $(p_1 - p_2) = D_0$, y una estimación agrupada para una p común ya no se usa en el error estándar. El estadístico de prueba modificado es

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - D_0}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

Aun cuando este estadístico de prueba no se usa con frecuencia, ¡el procedimiento no es diferente de otras pruebas de muestra grande que ya hemos aprendido!

5.5 PRUEBA DE HIPÓTESIS PARA UNA Y DOS VARIANZAS

La varianza como medida de dispersión es importante dado que nos ofrece una mejor visión de dispersión de datos. Por ejemplo: si se determina que la población califica en promedio con 6 el desempeño del gobierno; al decir que la varianza es de cero (y por lo tanto la desviación estándar es de cero) podemos confiar en que casi la misma calificación le asignaría toda la población. En otras palabras, en general la población en su conjunto evalúa al gobierno con la misma calificación, ya que no hay variación o dispersión en dicha calificación. Por el contrario, con la misma calificación promedio de 6 pero con una varianza muy alta podemos interpretar que hay gente contenta con el gobierno que le ha asignado calificaciones por arriba del 6; pero hay un conjunto poblacional muy molesto con el gobierno que asigna calificaciones por debajo del 6. Este tipo de información sólo es posible mediante el análisis de la varianza.

Otro campo del conocimiento donde la varianza se ocupa en gran medida es en control de calidad: cuando un producto se elabora el área de control de calidad busca que los productos estén dentro de ciertos límites de tolerancia, pero también que la variabilidad de un producto sea lo menor posible.

Se considera que la población sigue una distribución de probabilidad normal, para lo cual usamos el siguiente estadístico de prueba:

$$\chi^2 = \frac{(n-1)\sigma^2}{\sigma^2}$$

Este estadístico de prueba se conoce como chi cuadrada

$$gl = n - 1$$

La prueba estadística de la hipótesis nula

$$H_0 : \sigma_1^2 = \sigma_2^2$$

utiliza el estadístico de prueba

$$F = \frac{s_1^2}{s_2^2}$$

Cuando la hipótesis alternativa implica una prueba de una cola, esto es,

$$H_a : \sigma_1^2 > \sigma_2^2$$

se puede hallar el valor crítico de cola derecha para rechazar H_0 directamente de la tabla 6 del Apéndice. Sin embargo, cuando la hipótesis alternativa requiera una prueba de dos colas, es decir,

$$H_0 : \sigma_1^2 \neq \sigma_2^2$$

la región de rechazo se divide entre las colas superior e inferior de la distribución F. Estos valores críticos de cola izquierda *no se dan* en la tabla 6 por la siguiente razón: usted es libre de decidir a cuál de las dos poblaciones llamar "Población 1". Si siempre elige llamar "Población 1" a la población con la varianza muestral *más grande*, entonces el valor observado de su estadística de prueba siempre estará en la cola derecha de la distribución F. Aun cuando la mitad de la región de rechazo, el área $\alpha/2$ a su izquierda, estará en la cola inferior de la distribución, nunca será necesario usarla. Pero recuerde estos puntos, para una prueba de dos colas:

- El área de la cola derecha de la región de rechazo es sólo $\alpha/2$.
- El área a la derecha del estadístico de prueba observado es sólo (valor p)/2.

Los procedimientos formales para una prueba de hipótesis y un intervalo de confianza $(1 - \alpha)100\%$ para dos varianzas poblacionales se muestran a continuación.

PRUEBA DE HIPÓTESIS RESPECTO A LA IGUALDAD DE DOS VARIANZAS POBLACIONALES

1. Hipótesis nula: $H_0 : \sigma_1^2 = \sigma_2^2$
2. Hipótesis alternativa:

Prueba de una cola	**Prueba de dos colas**
$H_a : \sigma_1^2 > \sigma_2^2$ (o $H_a : \sigma_1^2 < \sigma_2^2$)	$H_a : \sigma_1^2 \neq \sigma_2^2$

3. Estadístico de prueba:

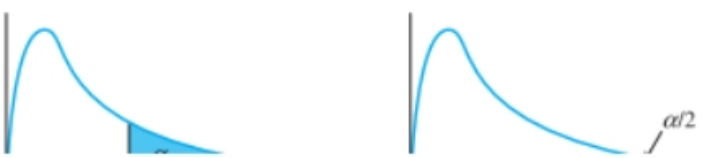
Prueba de una cola	**Prueba de dos colas**
$F = \dfrac{s_1^2}{s_2^2}$	$F = \dfrac{s_1^2}{s_2^2}$

   donde s_1^2 es la varianza muestral más grande.

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	**Prueba de dos colas**
$F > F_\alpha$	$F > F_{\alpha/2}$
o cuando valor $p < \alpha$	

Los valores críticos de F_α y $F_{\alpha/2}$ están basados en $gl_1 = (n_1 - 1)$ y $gl_2 = (n_2 - 1)$. Estos valores tabulados, para $a = .100, .050, .025, .010$ y $.005$, se pueden hallar usando la tabla 6 del Apéndice.



Suposiciones: Las muestras se seleccionan al azar y en forma independiente de las poblaciones normalmente distribuidas.

INTERVALO DE CONFIANZA PARA σ_1^2/σ_2^2

$$\left(\frac{s_1^2}{s_2^2}\right)\frac{1}{F_{gl_1,gl_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2}\right)F_{gl_2,gl_1}$$

donde $gl_1 = (n_1 - 1)$ y $gl_2 = (n_2 - 1)$. F_{gl_1,gl_2} es el valor crítico tabulado de F correspondiente a gl_1 y gl_2 grados de libertad en el numerador y denominador de F, respectivamente, con área $\alpha/2$ a su derecha.

Suposiciones: Las muestras se seleccionan al azar y en forma independiente de poblaciones normalmente distribuidas.

EJEMPLO 5.7

Un experimentador está preocupado porque la variabilidad de respuestas que usan dos procedimientos experimentales diferentes puede no ser igual. Antes de realizar su investigación, lleva a cabo un estudio previo con muestras aleatorias de 10 y 8 respuestas y obtiene $s_1^2 = 7.14$ y $s_2^2 = 3.21$, respectivamente. ¿Las varianzas muestrales presentan suficiente evidencia para indicar que las varianzas poblacionales son desiguales?

Solución Suponga que las poblaciones tienen distribuciones de probabilidad que son razonablemente en forma de montículo y por tanto satisfacen, para todos los fines prácticos, la suposición de que las poblaciones son normales. Se desea probar estas hipótesis:

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{contra} \quad H_a : \sigma_1^2 \neq \sigma_2^2$$

Usando la tabla 6 del Apéndice para $\alpha/2 = .025$, se puede rechazar H_0 cuando $F > 4.82$ con $\alpha = .05$. El valor calculado de la estadística de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{7.14}{3.21} = 2.22$$

Debido a que el estadístico de prueba no cae en la región de rechazo, no se puede rechazar H_0: $\sigma_1^2 = \sigma_2^2$. Por tanto, hay insuficiente evidencia para indicar una diferencia en las varianzas poblacionales.

EJEMPLO 5.8

Consulte el ejemplo 5.7 y encuentre un intervalo de confianza de 90% para σ_1^2/σ_2^2.

Solución El intervalo de confianza de 90% para σ_1^2/σ_2^2 es

$$\left(\frac{s_1^2}{s_2^2}\right)\frac{1}{F_{gl_1,gl_2}} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{s_1^2}{s_2^2}\right)F_{gl_2,gl_1}$$

donde

$$s_1^2 = 7.14 \qquad s_2^2 = 3.21$$
$$gl_1 = (n_1 - 1) = 9 \qquad gl_2 = (n_2 - 1) = 7$$
$$F_{9,7} = 3.68 \qquad F_{7,9} = 3.29$$

Sustituyendo estos valores en la fórmula para el intervalo de confianza, se obtiene

$$\left(\frac{7.14}{3.21}\right)\frac{1}{3.68} < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{7.14}{3.21}\right)3.29 \quad \text{o} \quad .60 < \frac{\sigma_1^2}{\sigma_2^2} < 7.32$$

La estimación calculada del intervalo de .60 a 7.32 incluye 1.0, el valor hipotético en H_0. Esto indica que es muy posible que $\sigma_1^2 = \sigma_2^2$ y por tanto concuerda con las conclusiones de prueba. No rechace $H_0 : \sigma_1^2 = \sigma_2^2$.

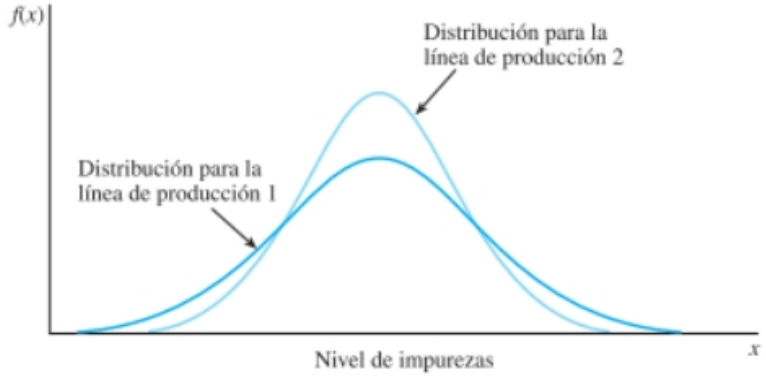
La función de *Excel* llamada **F.TEST** (**FTEST** en *Excel 2007* y versiones anteriores) puede usarse para realizar la prueba F para la igualdad de varianzas cuando ha introducido los datos sin procesar en la hoja de cálculo. El comando de *MINITAB* **Stat ▶ Basic Statistics ▶ 2 Variances** es un poco más flexible, porque permite introducir ya sea datos sin procesar o estadísticas de resumen para ejecutar la prueba F. Además, *MINITAB 16* calcula intervalos de confianza para la razón de dos varianzas o dos desviaciones estándar (lo cual todavía no se ha expuesto). La salida impresa de MINITAB 16 para el ejemplo 5.7, que contiene el estadístico F y su valor p, se muestra en la figura 5.4.

FIGURA 5.4
Salida de *MINITAB* para el ejemplo 5.7

Prueba y CI para dos varianzas

```
Null hypothesis            Sigma(1) / Sigma(2) = 1
Alternative hypothesis   Sigma(1) / Sigma(2) not = 1
Significance level       Alpha = 0.05

Statistics
Sample     N   StDev    Variance
1         10   2.672       7.140
2          8   1.792       3.210
Ratio of standard deviations = 1.491
Ratio of variances = 2.224


95% Confidence Intervals
CI for
Distribution    CI for StDev        Variance
of Data               Ratio              Ratio
Normal          (0.679, 3.055)   (0.461, 9.335)
Test
Method              DF1   DF2   Statistic   P-Value
F Test (normal)       9     7        2.22     0.304
```

FIGURA 5.5
Distribuciones de mediciones de impurezas para dos líneas de producción



EJEMPLO 5.9

La variabilidad en la cantidad de impurezas, presente en un lote de un producto químico empleado para un proceso particular, depende del tiempo en que el proceso está en operación. Un fabricante que utiliza dos líneas de producción, 1 y 2, ha hecho un ligero ajuste a la línea 2, esperando con ello reducir la variabilidad así como el promedio de la cantidad de impurezas en el producto químico. Muestras de $n_1 = 25$ y $n_2 = 25$ mediciones de los dos lotes dan estas medias y varianzas:

$$\bar{x}_1 = 3.2 \qquad s_1^2 = 1.04$$
$$\bar{x}_2 = 3.0 \qquad s_2^2 = .51$$

¿Los datos presentan suficiente evidencia para indicar que la variabilidad del proceso es menor para la línea 2?

Solución El experimentador piensa que el promedio de los niveles de impurezas son los mismos para las dos líneas de producción pero que su ajuste puede haber disminuido la variabilidad de los niveles para la línea 2, como se ilustra en la figura 5.5. Este ajuste sería bueno para la compañía porque disminuiría la probabilidad de hacer envíos del producto químico con niveles de impureza inaceptablemente altos.

Para determinar la disminución en variabilidad, la prueba de hipótesis es

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{contra} \quad H_a : \sigma_1^2 > \sigma_2^2$$

y el valor observado para este estadístico de prueba es

$$F = \frac{s_1^2}{s_2^2} = \frac{1.04}{.51} = 2.04$$

Usando el método del valor p, se puede limitar el valor p de una cola usando la tabla 6 del Apéndice con $gl_1 = gl_2 = (25 - 1) = 24$. El valor observado de F cae entre $F_{.050} = 1.98$ y $F_{.025} = 2.27$, de modo que $.025 <$ valor $p < .05$. Los resultados son juzgados como significativos al nivel del 5%, y H_0 es rechazada. Se concluye que la variabilidad de la línea 2 es menor que la de la línea 1.

La prueba F de la diferencia en dos varianzas poblacionales completa la batería de pruebas que usted ha aprendido en este capítulo para hacer inferencias acerca de parámetros poblacionales bajo estas condiciones:

- Los tamaños muestrales son pequeños.
- La muestra o muestras son sacados de poblaciones normales.

Usted encontrará que las distribuciones F y χ_2, así como la distribución t de Student, son muy importantes en otras aplicaciones en los capítulos que siguen. Se usarán para diferentes estimadores diseñados para contestar distintos tipos de preguntas inferenciales, pero las técnicas básicas para hacer inferencias siguen siendo iguales.

5.6 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

1. **a.** Defina α y β para una prueba estadística de hipótesis.

b. Para un tamaño muestral fijo n, si el valor de α disminuye, ¿cuál es el efecto en β?

c. Para disminuir α y β para un valor alternativo particular de μ, ¿cómo debe cambiar el tamaño muestral?

2. ¿Cuál es el valor p para una prueba de hipótesis? ¿Cómo se calcula para una prueba de muestra grande?

3. ¿Qué condiciones deben satisfacerse para que la prueba z pueda usarse para probar una hipótesis relativa a una media poblacional μ?

4. Defina la potencia de una prueba estadística. Conforme el valor alternativo de μ se aleja de μ_0, ¿cómo se afecta la potencia?

5. **Color de máquinas lavadoras** Un fabricante de máquinas lavadoras automáticas suministra un modelo particular en uno de tres colores. De las primeras mil lavadoras vendidas, se observa que

400 eran del primer color. ¿Puede usted concluir que más de un tercio de todos los compradores tienen preferencia por el primer color?

a. Encuentre el valor p para la prueba.

b. Si usted piensa realizar su prueba usando $\alpha = .05$, ¿cuáles serán las conclusiones de su prueba?

6. Pesca de robalo El factor pH es una medida de la acidez o alcalinidad del agua. Una lectura de 7.0 es neutral; valores de más de 7.0 indican alcalinidad y debajo de 7.0 implican acidez. Loren Hill establece que la mejor probabilidad de pescar robalo es cuando el pH del agua está entre 7.5 y 7.9.[1] Supongamos que usted sospecha que la lluvia ácida está bajando el pH de su lugar de pesca favorito y desea determinar si el pH es menor que 7.5.

a. Indique las hipótesis alternativa y nula que elegiría para una prueba estadística.

b. ¿La hipótesis alternativa de la parte **a** implica una prueba de una o de dos colas? Explique.

c. Suponga que una muestra aleatoria de 30 especímenes de agua dieron lecturas de pH con $\bar{x} = 7.3$ y $s = .2$. Con sólo un vistazo a los datos, ¿piensa usted que la diferencia $\bar{x} - 7.5 = -.2$ es suficientemente grande para indicar que la media de pH de las muestras de agua es menor que 7.5? (*No* realice la prueba.)

d. Ahora realice una prueba estadística de las hipótesis de la parte **a** e indique sus conclusiones. Pruebe usando $\alpha = .05$. Compare su decisión basada estadísticamente con su decisión intuitiva de la parte **c**.

7. Biomasa La biomasa para bosques tropicales, considerada de 35 kilogramos por metro cuadrado (kg/m^2), puede de hecho ser demasiado alta y que los valores de biomasa tropical varían de una región a otra, de unos 5 a 55 kg/m^2.[2] Supongamos que usted mide la biomasa tropical en 400 lugares de un metro cuadrado, seleccionados al azar y obtiene $\bar{x} = 31.75$ y $s = 10.5$. ¿Los datos presentan suficiente evidencia para indicar que los científicos valoran en exceso la media de biomasa para bosques tropicales y que la media es de hecho más baja que lo estimado?

a. Indique las hipótesis nula y alternativa a probar.

b. Localice la región de rechazo para la prueba con $\alpha = .01$.

c. Realice la prueba y exprese sus conclusiones.

8. Adolescentes y estrés social En un estudio para comparar diferencias étnicas de estrés social en adolescentes, unos investigadores reclutaron personas de tres escuelas de enseñanza media en Houston, Texas.[3] Una tabulación de respuestas de estudiantes a una pregunta relacionada con su condición socioeconómica (SES), comparada con otras familias en las que los estudiantes seleccionarían una de cinco respuestas (mucho peor, un poco peor, más o menos igual, mejor o mucho mejor) resultó en la tabulación que sigue.

	Euro-americanos	Afro-americanos	De origen hispano	De origen asiático
Tamaño muestral	144	66	77	19
Más o menos igual	68	42	48	8

a. ¿Estos datos apoyan la hipótesis de que la proporción de afroestadounidenses adolescentes, que dicen que su condición socioeconómica es "más o menos igual", es mejor que la de adolescentes estadounidenses de origen hispano?

b. Encuentre el valor p para la prueba.

c. Si usted planea probar usando $\alpha = .05$, ¿cuál es su conclusión?

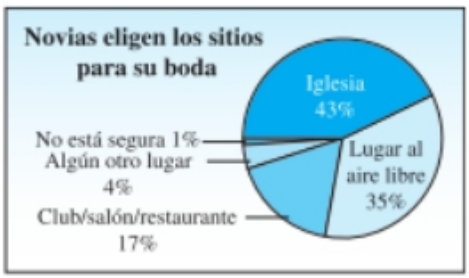
9. Bajar de peso En una comparación de la reducción media de peso en 1 mes, para mujeres de 20 a 30 años, estos datos muestrales se obtuvieron para cada una de dos dietas:

	Dieta I	Dieta II
Tamaño muestral n	40	40
Media muestral $\bar{x}$ (lb)	10	8
Varianza muestral s^2	4.3	5.7

¿Los datos dan suficiente evidencia para indicar que la dieta I produce una mayor reducción media de peso que la dieta II? Use $\alpha = .05$.

10. Experimento de un laberinto En un estudio de una carrera en un laberinto, una rata corre en un laberinto en forma de T y se registra el resultado de cada carrera. Una recompensa en forma de alimento se coloca siempre en la salida derecha. Si ocurre un aprendizaje, la rata elegirá la salida de la derecha con más frecuencia que la de la izquierda. Si no hay aprendizaje, la rata elegiría al azar cualquiera de las dos salidas. Suponga que a la rata se le dan $n = 100$ carreras en el laberinto y que ella elige la salida derecha $x = 64$ veces. ¿Concluiría usted que tiene lugar un aprendizaje? Use el método del valor p y tome una decisión con base en este valor p.

11. ¿Irá a la capilla? Si decide casarse, ¿qué tipo de sitio elegiría para la boda? Una instantánea (Snapshot) de *USA Today* afirma que 43% de todas las novias eligen una iglesia para su boda.[4]



Por Michelle Healey y Veronica Salazar, *USA TODAY*

Fuente: 'BRIDES' Magazine

En un estudio de seguimiento, en una muestra aleatoria de 100 novias se encontró que 46 de las participantes habían elegido o elegirían una iglesia para su boda. ¿Esta muestra contradice la cifra mencionada de 43%? Pruebe al nivel de significancia $\alpha = .05$.

12. Nadadores de estilo de pecho ¿Cuánto tiempo de entrenamiento es necesario para ser nadador de estilo de pecho de clase mundial? Un estudio publicado en *The American Journal of Sports Medicine* informó del número de metros por semana nadados por dos grupos de nadadores, los que compitieron sólo en estilo de pecho y los que compitieron en combinado individual (que incluye estilo de pecho). Se registró el número de metros por semana que ejercitan el nado de pecho y el resumen de las estadísticas se muestran a continuación.[5]

	Nado de pecho	Combinado individual
Tamaño muestral	130	80
Media muestral	9 017	5 853
Desviación estándar muestral	7 162	1 961

¿Hay suficiente evidencia para indicar una diferencia en el número promedio de metros nadados por estos dos grupos de nadadores? Pruebe usando $\alpha = .01$.

---

[1] Loren Hill, *Bassmaster*, septiembre/octubre de 1980.

[2] *Science News* 136 (19 de agosto de 1989):124.

[3] Jeeseung Choi, Janet Meininger y Robert E. Roberts, "Ethnic Differences in Adolescents' Mental Distress, Social Stress, and Resources", *Adolescence*, vol. 41, núm. 162, Verano de 2006, pp. 263-278.

[4] "Brides-to-be Pick Their Wedding Sites", http://www.usatoday.com/news/snapshot.htm, 6 de julio de 2010.

[5] Kurt Grote, T. L. Lincoln y J. G. Gamble, "Hip Adductor Injury in Competitive Swimmers", *The American Journal of Sports Medicine* 32, núm. 1 (enero/febrero de 2004):104.

EJERCICIOS

TÉCNICAS BÁSICAS

5.1 Encuentre las regiones de rechazo apropiadas para el estadístico de prueba z de muestras grandes en estos casos:

a. Una prueba de cola derecha con $\alpha = .01$.

b. Una prueba de dos colas al nivel de significancia de 5%.

5.2 Una muestra aleatoria de $n = 35$ observaciones de una población cuantitativa produjo una media de $\bar{x} = 2.4$ y una desviación estándar $s = .29$. Suponga que el objetivo de su investigación es demostrar que la media poblacional μ excede de 2.3.

a. Dé la hipótesis nula y alternativa para la prueba.

b. Localice la región de rechazo para la prueba usando un nivel de significancia de 5%.

c. Encuentre el error estándar de la media.

d. Antes de realizar la prueba, use su intuición para decidir si es probable la media muestral $\bar{x} = 2.4$, suponiendo que $\mu = 2.3$. Ahora realice la prueba. ¿Los datos dan suficiente evidencia para indicar que $\mu > 2.3$?

5.3 Una muestra aleatoria de $n = 1\,000$ observaciones de una población binomial produjo $x = 279$.

a. Si su hipótesis de investigación es que p sea menor que .3, ¿qué debe elegir para su hipótesis alternativa? ¿Y para su hipótesis nula?

b. ¿Cuál es el valor crítico que determina la región de rechazo para su prueba con $\alpha = .05$?

c. ¿Los datos dan suficiente evidencia para indicar que p es menor que .3? Use un nivel de significancia de 5%.

5.4 Muestras aleatorias independientes de dos poblaciones normales produjeron las varianzas dadas:

Tamaño muestral	Varianza muestral
13	18.3
13	7.9

a. ¿Los datos dan suficiente evidencia para indicar que $\sigma_1^2 > \sigma_2^2$? Pruebe usando $\alpha = .05$.

b. Encuentre el valor p aproximado para la prueba e interprete su valor.

APLICACIONES

5.5 Potencia de un antibiótico Un fabricante de medicamentos declaró que la potencia media de uno de sus antibióticos fue 80%. Se probó una muestra aleatoria de $n = 100$ cápsulas y produjo una media muestral de $\bar{x} = 79.7\%$ con una desviación estándar de $s = .8\%$. ¿Los datos presentan suficiente evidencia para refutar lo dicho por el fabricante? Sea $\alpha = .05$.

a. Exprese la hipótesis nula a ser probada.

b. Exprese la hipótesis alternativa.

c. Realice una prueba estadística de la hipótesis nula y exprese su conclusión.

5.6 ¿Cura para el resfriado común? Se planeó un experimento para comparar el tiempo medio (en días), necesario para recuperarse de un resfriado común en personas suministrándoles una dosis diaria de 4 miligramos (mg) de vitamina C en contraste con aquellas a quienes no se les dio un suplemento vitamínico. Suponga que se seleccionaron al azar 35 adultos para cada categoría del tratamiento y que los tiempos medios de recuperación y desviaciones estándar para los dos grupos fueron los siguientes:

	Sin suplemento vitamínico	4 mg de vitamina C
Tamaño muestral	35	35
Media muestral	6.9	5.8
Desviación estándar muestral	2.9	1.2

a. Suponga que el objetivo de su investigación es demostrar que el uso de vitamina C reduce el tiempo medio necesario para recuperarse de un resfriado común y sus complicaciones. Dé las hipótesis nula y alternativa para la prueba. ¿Esta prueba es de una o de dos colas?

b. Realice la prueba estadística de la hipótesis nula de la parte **a** y exprese su conclusión. Pruebe usando $\alpha = .05$.

5.7 Tarifas aéreas más baratas ¿Busca una tarifa aérea excelente? *Consumer Reports*[1] tiene varios consejos sobre cómo reducir sus costos, los cuales incluyen ser flexible respecto a las fechas y horarios de viaje, revisar múltiples sitios web y saber cuándo reservar su vuelo. Una sugerencia implicaba comprobar las tarifas en aeropuertos "secundarios" —aeropuertos que podrían estar un poco más lejos de su casa, pero donde las tarifas son más bajas. Por ejemplo, el promedio de todos los precios de boletos nacionales en el Aeropuerto Internacional de Los Ángeles (LAX) fue cotizado en $349 en comparación con un precio promedio de $287 en el cercano Aeropuerto Internacional de Ontario (ONT). Suponga que estas estimaciones se basaron en muestras aleatorias de 1 000 boletos nacionales en cada aeropuerto y que la desviación estándar de los precios en ambos aeropuertos fue $200.

a. ¿Hay suficiente evidencia para indicar que la media de los precios de los boletos difiere para estos dos aeropuertos en el nivel de significancia $\alpha = .05$? Use la prueba z de muestra grande. ¿Cuál es el valor p de esta prueba?

b. Construya un intervalo de confianza de 95% para $(\mu_1 - \mu_2)$. ¿Este intervalo confirma sus conclusiones de la parte **a**?

5.8 Obesidad infantil De acuerdo con una encuesta en la revista *PARADE*, casi la mitad de los padres dicen que el peso de sus hijos está bien.[2] Sólo 9% de los padres describe a sus hijos como obesos. No obstante, la Asociación Americana de Obesidad dice que el número de niños y adolescentes con sobrepeso es al menos de 15%. Suponga que su muestra es $n = 750$ padres y el número de quienes describen que sus hijos tienen sobrepeso es $x = 68$.

a. ¿Cómo procedería para probar la hipótesis de que la proporción de padres que describen a sus hijos como obesos es menor que la proporción real publicada por la Asociación Americana de Obesidad?

b. ¿Qué conclusión se puede sacar de estos datos al nivel de significancia de $\alpha = .05$?

c. ¿Cuál es el valor p asociado con esta prueba?

5.9 Genética de plantas Se cruzó una peonia con pétalos rojos con otra planta que tenía pétalos rayados. Un genetista declara

que 75% de los descendientes que resulten de esta cruza tendrán flores rojas. Para probar su dicho, se seleccionaron y germinaron 100 semillas de esta cruza, y 58 plantas tuvieron pétalos rojos.

a. ¿Qué hipótesis debería usar para probar la declaración del genetista?

b. Calcule el estadístico de prueba y su valor p. Use el valor p para evaluar la significancia estadística de los resultados al nivel del 1%.

5.10 Prueba de degustación En una prueba de degustación en igualdad de condiciones de alimentos de marca propia de supermercado contra marcas nacionales, *Consumer Reports* encontró que era difícil hallar una diferencia de sabor entre las dos.[3] Si la marca nacional en efecto es mejor que la marca propia, debería declararse la mejor más de 50% de las veces.

a. Formule las hipótesis nula y alternativa que serán probadas. ¿Ésta es una prueba de una cola o de dos?

b. Suponga que, de las 35 categorías de alimentos usadas para la prueba de degustación, se encontró que la marca nacional fue mejor que la propia en ocho de las comparaciones de sabor. Use esta información para probar las hipótesis de la parte **a**. Use $\alpha = .01$. ¿Qué conclusiones prácticas obtiene de los resultados?

5.11 El mejor amigo del hombre La Sociedad Protectora de Animales informa que en Estados Unidos hay alrededor de 77.5 millones de perros con dueño y que aproximadamente 40% de todas las familias tienen al menos un perro.[4] En una muestra aleatoria de 300 familias, 114 dijeron que tenían al menos un perro. ¿Estos datos dan suficiente evidencia para indicar que la proporción de hogares con al menos un perro es diferente de la publicada por la Sociedad Protectora de Animales? Pruebe usando $\alpha = .05$.

5.12 Baterías de litio La estabilidad de las mediciones en un producto manufacturado es importante para mantener su calidad. Un fabricante de baterías de litio, como las que se usan para cámaras digitales, sospecha que una de las líneas de producción producía baterías con una duración que variaba de manera amplia. Para probar esta teoría, seleccionó al azar $n = 50$ baterías de la línea sospechosa y $n = 50$ de una línea que se evaluó como "en control". Luego midió la duración (en horas) hasta su reducción a 0.85 V con una carga de 5 ohmios para ambas muestras. Las medias y varianzas muestrales para las dos muestras fueron las siguientes:

Línea sospechosa	Línea "en control"
$\bar{x}_1 = 9.40$	$\bar{x}_2 = 9.25$
$s_1 = .25$	$s_2 = .12$

a. ¿Los datos proporcionan evidencia suficiente para indicar que las baterías producidas por la "línea sospechosa" tienen una varianza más grande en la duración que aquellas producidas por la línea que se supone está en control? Pruebe usando $\alpha = .05$.

b. Encuentre el valor p aproximado para la prueba e interprete su valor.

c. Construya un intervalo de confianza del 90% para la razón de la varianza.

5.13 Impurezas Un fabricante farmacéutico compra un material particular a dos proveedores diferentes. El nivel medio de impurezas en la materia prima es aproximadamente igual para ambos proveedores, pero el fabricante está preocupado por la variabilidad de las impurezas de un embarque a otro. Para comparar la variación en el porcentaje de impurezas para los dos proveedores, el fabricante selecciona 10 envíos de cada uno de ellos y mide el porcentaje de impurezas en la materia prima para cada embarque. Las medias y varianzas muestrales se presentan en la tabla.

Proveedor A	Proveedor B
$\bar{x}_1 = 1.89$	$\bar{x}_2 = 1.85$
$s_1^2 = .273$	$s_2^2 = .094$
$n_1 = 10$	$n_2 = 10$

a. ¿Los datos dan suficiente evidencia para indicar una diferencia en la variabilidad de los niveles de impureza de envíos para los dos proveedores? Pruebe usando $\alpha = .01$. Con base en los resultados de su prueba, ¿qué recomendación se haría al fabricante de productos farmacéuticos?

b. Encuentre un intervalo de confianza de 99% para σ_2^2 e interprete sus resultados.

Fuentes

1. "8 Ways to Land a Great Airfare," *Consumer Reports*, junio de 2010, y www.consumerreports.org/cro/magazine-archive/2010/june/money/airfares/overview/index.htm, 2 de septiembre de 2010.
2. Dianne Hales, "We're Changing the Way We Eat," *PARADE*, 12 de noviembre de 2006, pp. 4–5.
3. "It Pays to Buy Store Brands," *Consumer Reports*, octubre de 2009 y www.consumerreports.org/cro/magazine-archive/october-2009/shopping/buying-storebrands/overview/buying-storebrands-ov.htm, 4 de septiembre de 2010.
4. "U.S. Pet Ownership Statistics," http://www.humanesociety.org/issues/pet_overpopulation/facts/pet_ownership_statistics.html, 3 de septiembre de 2010.

6

Análisis de varianza

© chirapbogdan/Shutterstock.com

OBJETIVOS GENERALES

La cantidad de información contenida en una muestra es afectada por diversos factores que el experimentador puede o no controlar. Este capítulo presenta tres *diseños experimentales* diferentes, dos de los cuales son extensiones directas de los diseños pareados y no pareados del capítulo 4. Una nueva técnica se utiliza para determinar el modo en que diferentes factores experimentales afectan la respuesta promedio.

ÍNDICE DEL CAPÍTULO

- Análisis de varianza
- Diseño completamente aleatorizado
- Diseño por bloques aleatorizados
- Método de Tukey de comparaciones pareadas

NECESITO SABER...

Cómo determinar si mis cálculos son precisos

6.1 DEFINICIÓN, CONCEPTOS FUNDAMENTALES

La forma en que se selecciona una muestra se denomina *plan de muestreo* o *diseño experimental* y determina la cantidad de información en la muestra. Algunas investigaciones comprenden un **estudio observacional**, en el que el investigador no produce en realidad los datos sino que sólo *observa* las características de los datos que ya existen. Casi todos los estudios muestrales, en donde se reúne información con un cuestionario, se encuentran en esta categoría. El investigador diseña un plan para recolectar los datos, llamado *plan de muestreo*, y a continuación utiliza los procedimientos estadísticos apropiados para obtener conclusiones acerca de la población o poblaciones de donde proviene la muestra.

Otra investigación comprende la **experimentación**. El investigador puede deliberadamente establecer una o más condiciones en las unidades experimentales para determinar su efecto en la respuesta. Veamos ahora algunos términos nuevos que usaremos para discutir el diseño de un experimento estadístico.

Definición Una **unidad experimental** es el objeto en el que se toma una medición (o mediciones).
Un **factor** es una variable independiente cuyos valores son controlados y variados por el experimentador.
Un **nivel** es el escenario de intensidad de un factor.
Un **tratamiento** es una combinación específica de niveles de factor.
La **respuesta** es la variable que es medida por el experimentador.

EJEMPLO 6.1 Un conjunto de personas se divide al azar en un grupo experimental y uno de control. Al grupo de control se le aplica un examen de aptitud después de haber tomado un desayuno completo; al grupo experimental se le aplica el mismo examen sin haber tomado desayuno. ¿Cuáles son los factores, niveles y tratamientos en este experimento?

Solución Las *unidades experimentales* son las personas en las que la *respuesta* (calificación de examen) se mide. El *factor* de interés podría describirse como "comida" y tiene dos *niveles*: "desayuno" y "sin desayuno". Como éste es el único factor controlado por el experimentador, los dos niveles de "desayuno" y "sin desayuno" también representan los *tratamientos* de interés en el experimento.

EJEMPLO 6.2 Suponga que el experimentador del ejemplo 6.1 empezó por seleccionar al azar 20 hombres y 20 mujeres para el experimento. Estos dos grupos se dividieron entonces al azar en 10 cada uno para los grupos experimental y de control. ¿Cuáles son los factores, niveles y tratamientos en este experimento?

Solución Ahora hay dos *factores* de interés para el experimentador y cada factor tiene dos *niveles*:

- "Género" a dos niveles: hombres y mujeres
- "Comida" a dos niveles: desayuno y sin desayuno

En este experimento más complejo hay cuatro *tratamientos*, uno para cada combinación específica de niveles de factor: hombres sin desayuno, hombres con desayuno, mujeres sin desayuno y mujeres con desayuno.

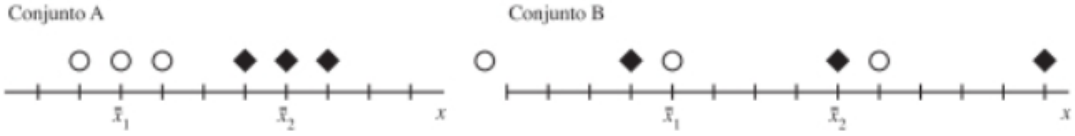
En este capítulo nos concentraremos en los experimentos que se han diseñado en tres formas diferentes y usaremos una técnica llamada *análisis de varianza* para juzgar los efectos de varios factores en la respuesta experimental. Dos de estos *diseños experimentales* son extensiones de los diseños pareados y no pareados del capítulo 4.

¿Qué es un análisis de varianza?

Las respuestas que se generan en una situación experimental siempre muestran cierta cantidad de *variabilidad*. En un **análisis de varianza** se divide la variación total de las mediciones de respuesta en partes que pueden ser atribuidas a varios *factores* de interés para el experimentador. Si el experimento se ha diseñado en forma apropiada estas partes pueden usarse entonces para contestar preguntas acerca de los efectos de los diversos factores en la respuesta de interés.

Se puede entender mejor la lógica subyacente en un análisis de varianza al analizar un experimento sencillo. Considere dos conjuntos de muestras seleccionadas al azar de las poblaciones 1 (◆) y 2 (○), cada uno con pares idénticos de medias; $\bar{x}_1$ y $\bar{x}_2$. Los dos conjuntos se observan en la figura 6.1. ¿Es más fácil detectar la diferencia en las dos medias cuando se vea el conjunto A o el conjunto B? Es probable que esté de acuerdo en que el conjunto A muestra la diferencia con mayor claridad. En el conjunto A, la variabilidad de las mediciones *dentro* de los grupos (los ◆ y los ○) es mucho menor que la variabilidad *entre* los dos grupos. En el conjunto B, hay más variabilidad *dentro* de los grupos (los ◆ y los ○), causando que los dos grupos se "mezclen" y hagan más difícil ver la diferencia *idéntica* en las medias.

FIGURA 6.1
Dos conjuntos de muestras con las mismas medias



La comparación que usted ha hecho intuitivamente es formalizada por el análisis de varianza. Es más, el análisis de varianza se puede usar no sólo para comparar dos medias sino también para hacer comparaciones de *más de dos* medias poblacionales y para determinar los efectos de varios factores en diseños experimentales más complejos. El análisis de varianza se apoya en estadísticas con distribuciones muestrales que son modeladas por la distribución *F*.

Suposiciones para un análisis de varianza

Las suposiciones necesarias para un análisis de varianza son semejantes a las requeridas para los estadísticos *t* de Student del capítulo 4. Cualquiera que sea el diseño experimental empleado para generar los datos, se debe suponer que las observaciones dentro de cada grupo de tratamiento están **normalmente distribuidas** con una **varianza común** σ^2. Al igual que en el capítulo 4, los procedimientos de análisis de varianza son bastante **robustos** cuando los tamaños muestrales son iguales y cuando los datos son de forma de montículo. Violar la suposición de una varianza común es más serio, en especial cuando los tamaños muestrales no son cercanamente iguales.

SUPOSICIONES PARA LA PRUEBA DE ANÁLISIS DE VARIANZA Y LOS PROCEDIMIENTOS DE ESTIMACIÓN

- Las observaciones dentro de cada población están distribuidas normalmente con una varianza común σ^2.
- Las suposiciones respecto al procedimiento de muestreo se especifican para cada diseño en las secciones que siguen.

Este capítulo describe el análisis de varianza para tres diseños experimentales diferentes. El primer diseño se basa en el muestreo aleatorio independiente de varias poblaciones y es una extensión de la *prueba t no pareada* del capítulo 4. El segundo es una extensión del diseño de *diferencia pareada* o *pares acoplados* y comprende una asignación aleatoria de tratamientos

dentro de conjuntos pareados de observaciones. El tercero es un diseño que permite juzgar el efecto de dos factores experimentales en la respuesta. Los procedimientos de muestreo necesarios para cada diseño se vuelven a expresar en sus secciones respectivas.

6.2 ANÁLISIS DE VARIANZA DE UNA VÍA O FACTOR

Uno de los diseños experimentales más sencillos es el **diseño completamente aleatorizado**, en el que muestras aleatorias se seleccionan de manera independiente de cada una de k poblaciones. Este diseño comprende sólo un *factor*, la población de donde proviene la medición, de aquí la designación como una **clasificación en una dirección**. Hay k *niveles* diferentes correspondientes a las k poblaciones, que también son los *tratamientos* para esta clasificación de una dirección. ¿Las k medias poblacionales son todas iguales, o al menos una media es diferente de las otras?

¿Por qué se necesita un nuevo procedimiento, el *análisis de varianza*, para comparar las medias poblacionales cuando ya se tiene disponible la prueba t de Student? Al comparar $k = 3$ medias se podría probar cada uno de los tres pares de hipótesis:

$$H_0 : \mu_1 = \mu_2 \qquad H_0 : \mu_1 = \mu_3 \qquad H_0 : \mu_2 = \mu_3$$

para averiguar dónde están las diferencias. No obstante, es preciso recordar que cada prueba que se realice está sujeta a la posibilidad de error. Para comparar $k = 4$ medias se necesitarían seis pruebas y se necesitarían 10 pruebas para comparar $k = 5$ medias. Cuantas más pruebas se realicen en un conjunto de mediciones más probable será que al menos una de las conclusiones sea incorrecta. El procedimiento de análisis de varianza provee una prueba general para juzgar la igualdad de las k medias poblacionales. Una vez que haya determinado si hay *en realidad* una diferencia en las medias se puede usar otro procedimiento para averiguar dónde están las diferencias.

¿Cómo se pueden seleccionar estas k muestras aleatorias? A veces las poblaciones existen en realidad y se puede usar un generador computarizado de números aleatorios o una tabla de números aleatorios para seleccionar al azar las muestras. Por ejemplo, en un estudio para comparar los tamaños promedio de reclamaciones de seguro médico en cuatro estados diferentes se podría usar una base de datos computarizada, proporcionada por las compañías de seguros médicos para seleccionar muestras aleatorias de los cuatro estados. En otras situaciones las poblaciones pueden ser *hipotéticas* y se pueden generar respuestas sólo después de que se han aplicado los tratamientos experimentales.

EJEMPLO 6.3 Un investigador está interesado en los efectos de cinco tipos de insecticidas a usar para controlar el gorgojo del algodón en campos algodoneros. Explique cómo poner en práctica un diseño completamente aleatorizado para investigar los efectos de los cinco insecticidas en la producción de cosechas.

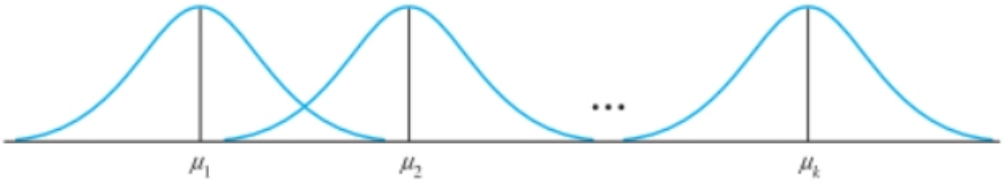
Solución La única forma de generar el equivalente de cinco muestras aleatorias de las poblaciones hipotéticas, correspondientes a los cinco insecticidas, es usar un método llamado **asignación aleatorizada**. Se escoge un número fijo de plantas de algodón para tratamiento y a cada una se le asigna un número aleatorio. Suponga que cada muestra debe tener un número igual de mediciones. Con el uso de un medio de aleatorización, se pueden asignar las primeras n_1 plantas escogidas para recibir el insecticida 1, las segundas n_2 plantas para recibir el insecticida 2 y así sucesivamente, hasta que se hayan asignado los cinco tratamientos.

Ya sea por *selección aleatoria* o *asignación aleatoria*, estos dos ejemplos resultan en un diseño completamente aleatorizado o clasificación en una dirección, para el cual se usa el análisis de varianza.

Prueba de hipótesis

Suponga que se desea comparar k medias poblacionales, $\mu_1, \mu_2, \ldots, \mu_k$, con base en muestras aleatorias independientes de tamaño $n_1, n_2, \ldots, n_k$ de poblaciones normales con una varianza común σ^2. Esto es, cada una de las poblaciones normales tiene la misma forma, pero sus ubicaciones podrían ser diferentes, como se ve en la figura 6.2.

FIGURA 6.2
Poblaciones normales con una varianza común pero medias diferentes



División de la variación total en un experimento

Sea x_{ij} la j–ésima medición ($j = 1, 2, \ldots, n_i$) en la i–ésima muestra. El procedimiento de análisis de varianza empieza por considerar la variación total en el experimento, que es medida por una cantidad llamada **suma total de cuadrados (TSS)**:

$$\text{TSS} = \Sigma(x_{ij} - \bar{x})^2 = \Sigma x_{ij}^2 - \frac{(\Sigma x_{ij})^2}{n}$$

Éste es el conocido numerador de la fórmula para la varianza muestral para todo el conjunto de $n_1 + n_2 + \ldots + n_k$ mediciones. La segunda parte de la fórmula de cálculo se denomina a veces **corrección para la media (CM)**. Si con G representamos el *gran total* de todas las n observaciones, entonces

$$\text{CM} = \frac{(\Sigma x_{ij})^2}{n} = \frac{G^2}{n}$$

Esta suma total de cuadrados (TSS) se divide en dos componentes. El primero, llamado **suma de cuadrados para tratamientos (SST)**, mide la variación entre las k medias muestrales:

$$\text{SST} = \Sigma n_i(\bar{x}_i - \bar{x})^2 = \Sigma\frac{T_i^2}{n_i} - \text{CM}$$

donde T_i es el total de observaciones para el tratamiento i. El segundo componente, llamado **suma de cuadrados para el error (SSE)**, se usa para medir la variación agrupada dentro de las k muestras:

$$\text{SSE} = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2$$

Esta fórmula es una extensión directa del numerador en la fórmula para la estimación agrupada de σ^2 del capítulo 4. Podemos demostrar algebraicamente que, en el análisis de varianza,

$$\text{TSS} = \text{SST} + \text{SSE}$$

Por tanto, es necesario calcular sólo dos de las tres sumas de cuadrados: TSS, SST y SSE, y la tercera se puede hallar por sustracción.

Cada una de las fuentes de variación, cuando es dividida entre sus apropiados **grados de libertad**, da una estimación de la variación en el experimento. Como TSS involucra n observaciones cuadradas, sus grados de libertad son $gl = (n - 1)$. Del mismo modo, la suma de cuadrados para tratamientos comprende k observaciones cuadradas y sus grados de libertad son $gl = (k - 1)$. Por último, la suma de cuadrados de error, una extensión directa de la estimación agrupada del capítulo 4, tiene

$$gl = (n_1 - 1) + (n_2 - 1) + \ldots + (n_k - 1) = n - k$$

Observe que los grados de libertad para tratamientos y error son aditivos, es decir,

$$gl\ (\text{total}) = gl\ (\text{tratamientos}) + gl\ (\text{error})$$

Estas dos fuentes de variación y sus respectivos grados de libertad se combinan para formar los **cuadráticos medios** como MS = SS/*gl*. La variación total en el experimento se exhibe entonces en una **tabla de análisis de varianza** (o **ANOVA**).

TABLA ANOVA PARA *k* MUESTRAS ALEATORIAS INDEPENDIENTES: DISEÑO COMPLETAMENTE ALEATORIZADO

Fuente	*gl*	SS	MS	*F*
Tratamientos	$k-1$	SST	MST = SST/$(k-1)$	MST/MSE
Error	$n-k$	SSE	MSE = SSE/$(n-k)$	
Total	$n-1$	TSS		

TIP MI CONSEJO

La columna marcada "SS" satisface:
TSS = SST + SSE

donde

$$\text{TSS} = \Sigma x_{ij}^2 - \text{CM}$$
$$= (\text{Suma de cuadrados de todos los valores } x) - \text{CM}$$

con

$$\text{CM} = \frac{(\Sigma x_{ij})^2}{n} = \frac{G^2}{n}$$

$$\text{SST} = \Sigma \frac{T_i^2}{n_i} - \text{CM} \qquad \text{MST} = \frac{\text{SST}}{k-1}$$

$$\text{SSE} = \text{TSS} - \text{SST} \qquad \text{MSE} = \frac{\text{SSE}}{n-k}$$

TIP MI CONSEJO

La columna marcada "*df*" siempre asciende a $n-1$

y

G = Gran total de todas las n observaciones
T_i = Total de todas las observaciones en la muestra i
n_i = Número de observaciones en la muestra i
$n = n_1 + n_2 + \ldots + n_k$

EJEMPLO 6.4

En un experimento para determinar el efecto de la nutrición sobre los intervalos de atención de los estudiantes de las escuelas básicas, un grupo de 15 estudiantes se asignó al azar a uno de tres planes de comidas: sin desayuno, desayuno ligero y desayuno completo. Sus intervalos de atención (en minutos) se registraron durante un periodo de lectura por la mañana y se muestran en la tabla 6.1. Construya la tabla de análisis de varianza para este experimento.

TABLA 6.1 Intervalos de atención de los estudiantes después de tres planes de comidas

Sin desayuno	Desayuno ligero	Desayuno completo
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12
$T_1 = 47$	$T_2 = 70$	$T_3 = 65$

Solución Para usar las fórmulas de cálculo, se necesitan los $k = 3$ tratamientos totales con $n_1 = n_2 = n_3 = 5$, $n = 15$ y $\Sigma x_{ij} = 182$. Entonces

$$\text{CM} = \frac{(182)^2}{15} = 2208.2667$$

$\text{TSS} = (8^2 + 7^2 + \ldots + 12^2) - \text{CM} = 2338 - 2208.2667 = 129.7333$ con $(n - 1) = (15 - 1) = 14$ grados de libertad,

$$\text{SST} = \frac{47^2 + 70^2 + 65^2}{5} - \text{CM} = 2266.8 - 2208.2667 = 58.5333$$

con $(k - 1) = (3 - 1) = 2$ grados de libertad, y por sustracción,

$$\text{SSE} = \text{TSS} - \text{SST} = 129.7333 - 58.5333 = 71.2$$

con $(n - k) = (15 - 3) = 12$ grados de libertad. Estas tres fuentes de variación, sus grados de libertad, sumas de cuadrados y cuadráticos medios se muestran en el área sombreada de las tablas ANOVA generadas por *MS Excel* y *MINITAB* y se dan en la figura 6.3. Se encontrarán instrucciones para generar esta salida impresa en la sección 6.4.

FIGURA 6.3a
Salida *MINITAB* para el ejemplo 6.4

ANOVA de una dirección: intervalo contra comida

Source	DF	SS	MS	F	P
Meal	2	58.53	29.27	4.93	0.027
Error	12	71.20	5.93		
Total	14	129.73			

S = 2.436 R-Sq = 45.12% R-Sq(adj) = 35.97%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
1	5	9.400	2.302
2	5	14.000	2.550
3	5	13.000	2.449

```
--+---------+---------+---------+-------
(---------*---------)
                     (---------*---------)
                  (---------*---------)
--+---------+---------+---------+-------
7.5       10.0      12.5      15.0
```

Pooled StDev = 2.436

FIGURA 6.3b
Salida de *MS Excel* para el ejemplo 6.4

Anova: factor único

RESUMEN

Grupos	Conteo	Suma	Promedio	Varianza
Sin	5	47	9.4	5.3
Ligero	5	70	14	6.5
Completo	5	65	13	6

ANOVA

Fuente de variación	SS	df	MS	F	Valor P	F crít.
Entre grupos	58.533	2	29.267	4.933	0.027	3.885
Dentro de grupos	71.2	12	5.933			
Total	129.733	14				

Las salidas de computadora dan alguna información adicional sobre la variación en el experimento. La sección inferior en *MINITAB* y la sección superior en *MS Excel* muestran las medias y desviaciones estándar (o varianzas) para los tres planes de comidas. Más importante aún es que se pueden ver en la sección superior en *MINITAB* y en la sección inferior en *MS Excel* dos columnas marcadas "F" y "P" ("valor F" y "valor P" en *Excel*). Podemos usar estos valores para probar una hipótesis respecto a la igualdad de las tres medias de tratamiento.

Prueba de igualdad de las medias de tratamiento

MI CONSEJO
MS = SS/*df*

Los *cuadráticos medios* en la tabla de análisis de varianza se pueden usar para probar la hipótesis nula

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k$$

contra la hipótesis alternativa

H_a: Al menos una de las medias es diferente de las otras

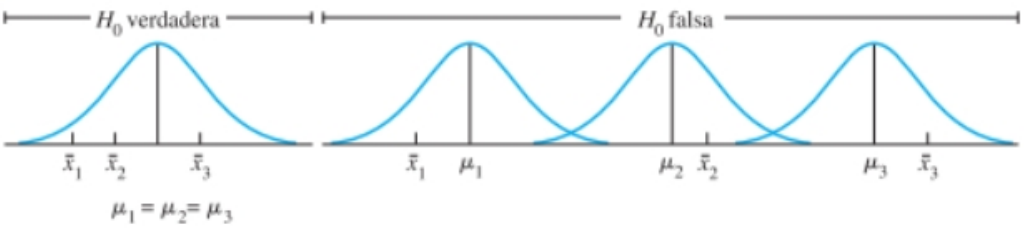
usando el siguiente argumento teórico:

- Recuerde que σ^2 es la varianza común para todas las poblaciones k. La cantidad

$$\text{MSE} = \frac{\text{SSE}}{n - k}$$

es una estimación agrupada de σ^2, un promedio ponderado de todas las k varianzas muestrales, sea H_0 verdadera o no.
- Si H_0 es verdadera, entonces la variación en las medias muestrales, medida por MST $= [\text{SST}/(k - 1)]$ también da una estimación insesgada de σ^2. No obstante, si H_0 es falsa y las medias poblacionales son diferentes, entonces MST, que mide la variación en las medias muestrales, será inusualmente *grande*, como se ve en la figura 6.4.

FIGURA 6.4
Medias muestrales sacadas de poblaciones idénticas contra diferentes



Variación entre grupos y dentro de los grupos

La siguiente pregunta obvia que se podría hacer se refiere a la naturaleza de las diferencias de las medias poblacionales. ¿Cuáles medias son diferentes de las otras? ¿Cómo se puede estimar la diferencia o posiblemente las medias individuales para cada uno de los tres tratamientos? Más adelante presentaremos un procedimiento que se puede usar para comparar todos los posibles pares de medias de tratamiento simultáneamente. No obstante, si hay un interés especial en una media particular o par de medias, se pueden construir intervalos de confianza usando los procedimientos de muestra pequeña del capítulo 4, con base en la distribución t de Student. Para una sola media poblacional, μ_i, el intervalo de confianza es

$$\bar{x}_i \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n_i}}\right)$$

donde $\bar{x}_i$ es la media muestral para el i–ésimo tratamiento. Del mismo modo, para una comparación de dos medias poblacionales, por ejemplo μ_i y μ_j, el intervalo de confianza es

$$(\bar{x}_i - \bar{x}_j) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Antes de que se puedan usar estos intervalos de confianza, sin embargo, persisten dos preguntas:

- ¿Cómo se calcula s o s^2, la mejor estimación de la varianza común σ^2?
- ¿Cuántos grados de libertad se usan para el valor crítico de t?

Para contestar estas preguntas, recuerde que en un análisis de varianza, el cuadrático medio de error, MSE, siempre da un estimador insesgado de σ^2 y usa información de todo el conjunto de mediciones. En consecuencia, es el mejor estimador disponible de σ^2, cualquiera que sea el procedimiento de prueba o estimación que se use. *Siempre* se debe usar

$$s^2 = \text{MSE} \qquad \text{con } gl = (n - k)$$

para estimar σ^2. Se puede hallar la raíz cuadrada positiva de este estimador, $s = \sqrt{\text{MSE}}$, en el último renglón de la figura 6.3a marcado "Pooled StDev".

DISEÑO COMPLETAMENTE ALEATORIZADO: INTERVALOS DE CONFIANZA $(1 - \alpha)100\%$ PARA UNA SOLA MEDIA DE TRATAMIENTO Y LA DIFERENCIA ENTRE DOS MEDIAS DE TRATAMIENTO

TIP MI CONSEJO

Los grados de libertad para intervalos de confianza son los *df* de **error**

Una sola media de tratamiento:

$$\bar{x}_i \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n_i}}\right)$$

Diferencia entre dos medias de tratamiento:

$$(\bar{x}_i - \bar{x}_j) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

con

$$s = \sqrt{s^2} = \sqrt{\text{MSE}} = \sqrt{\frac{\text{SSE}}{n - k}}$$

donde $n = n_1 + n_2 + \ldots + n_k$ y $t_{\alpha/2}$ está basada en $(n - k)$ *gl*.

EJEMPLO 6.5

El investigador del ejemplo 6.4 cree que los estudiantes que no toman desayuno tendrán intervalos de atención significativamente más cortos, pero que puede no haber diferencia entre aquellos que toman un desayuno ligero o un desayuno completo. Encuentre un intervalo de confianza de 95% para el promedio de intervalo de atención para estudiantes que no toman desayuno, así como un intervalo de confianza para la diferencia en los intervalos de atención promedio para quienes toman desayuno ligero contra los de desayuno completo.

Solución Para $s^2 = \text{MSE} = 5.9333$, para que $s = \sqrt{5.9333} = 2.436$ con $gl = (n - k) = 12$, se pueden calcular los dos intervalos de confianza:

- Para sin desayuno:

$$\bar{x}_1 \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n_1}}\right)$$

$$9.4 \pm 2.179\left(\frac{2.436}{\sqrt{5}}\right)$$

$$9.4 \pm 2.37$$

esto es, entre 7.03 y 11.77 minutos.

- Para los de desayuno ligero contra desayuno completo:

$$(\bar{x}_2 - \bar{x}_3) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_2} + \frac{1}{n_3}\right)}$$

$$(14 - 13) \pm 2.179\sqrt{5.9333\left(\frac{1}{5} + \frac{1}{5}\right)}$$

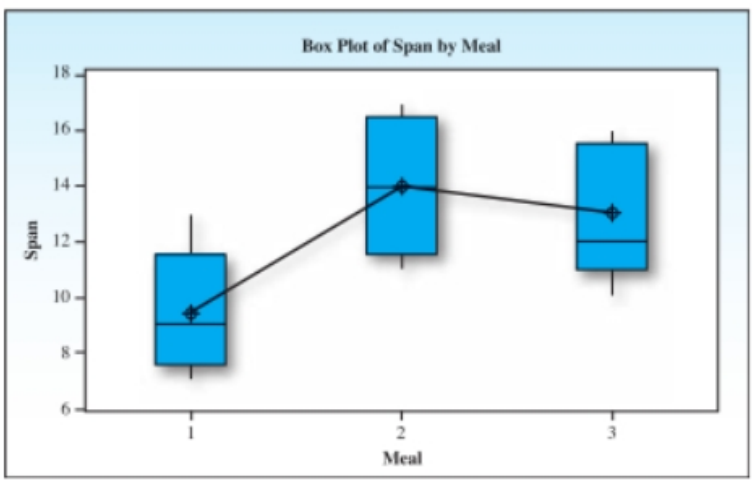
$$1 \pm 3.36$$

una diferencia de entre −2.36 y 4.36 minutos.

Se puede ver que el segundo intervalo de confianza no indica una diferencia en el promedio de intervalos de atención, para los estudiantes que tomaron un desayuno ligero contra los de desayuno completo, como sospechaba el investigador. Si el investigador, debido a creencias previas, desea probar los otros dos posibles pares de medias, es decir, sin desayuno contra desayuno ligero y sin desayuno contra desayuno completo, los métodos que se proporcionan en el siguiente apartado deben usarse para probar los tres pares.

Algunos programas de computadora tienen opciones de gráficas que dan una excelente descripción visual de datos y las k medias de tratamiento. Una de estas opciones en el programa *MINITAB* se ve en la figura 6.5. Las medias de tratamiento están indicadas por el símbolo ⊕ y están conectadas con rectas. Observe que la media "sin desayuno" parece ser un poco diferente de las otras dos medias, como sospechaba el investigador, aun cuando hay un pequeño traslape en las gráficas de caja. En el siguiente apartado presentamos un procedimiento formal para probar la significancia de las diferencias entre todos los pares de medias de tratamiento.

FIGURA 6.5
Gráficas de caja para el ejemplo 6.5



La distribución *F*

La prueba estadística

$$F = \frac{\text{MST}}{\text{MSE}}$$

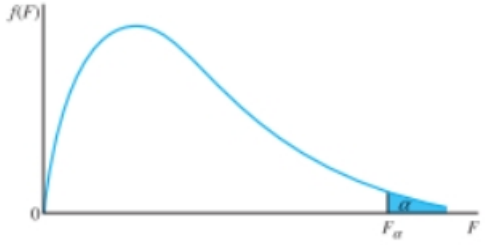
tiende a ser más grande que lo normal si H_0 es falsa. En consecuencia, se puede rechazar H_0 para valores grandes de F, usando una prueba estadística de *cola derecha*. Cuando H_0 es verdadera, esta prueba estadística tiene una distribución F con $gl_1 = (k - 1)$ y $gl_2 = (n - k)$ grados de libertad y se pueden usar valores críticos de *cola derecha* de la distribución F (de la tabla 6 del Apéndice) o valores p generados por computadora, para sacar conclusiones estadísticas acerca de la igualdad de las medias poblacionales.

TIP MI CONSEJO

Las pruebas *F* para tablas ANOVA son **siempre** de cola superior (derecha)

PRUEBA *F* PARA COMPARAR *K* MEDIAS POBLACIONALES

1. Hipótesis nula: H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$
2. Hipótesis alternativa: H_a: Uno o más pares de medias poblacionales difieren
3. Prueba estadística: $F = \text{MST/MSE}$, donde F está basada en $gl_1 = (k - 1)$ y $gl_2 = (n - k)$
4. Región de rechazo: rechazar H_0 si $F > F_\alpha$, donde F_α se encuentra en la cola superior de la distribución F (con $gl_1 = k - 1$ y $gl_2 = n - k$) o si el valor $p < \alpha$.



- Las muestras son seleccionadas al azar y en forma independiente de sus respectivas poblaciones.
- Las poblaciones están normalmente distribuidas con $\mu_1, \mu_2, \ldots, \mu_k$ y varianzas iguales, $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = \sigma^2$.

EJEMPLO 6.6

¿Los datos del ejemplo 6.4 dan suficiente evidencia para indicar una diferencia en el promedio de intervalos de atención, dependiendo del tipo de desayuno tomado por el estudiante?

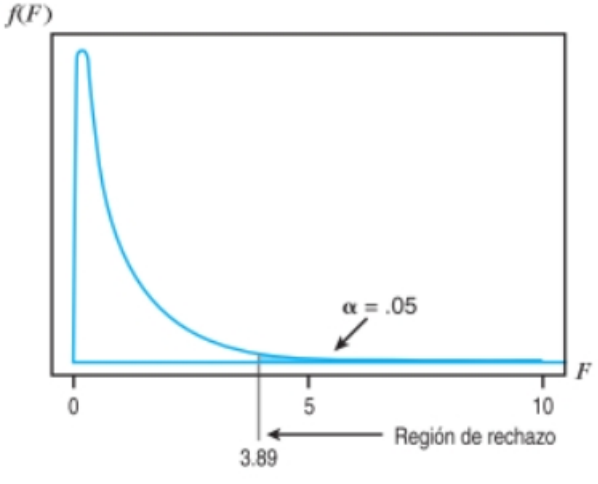
Solución Para probar H_0: $\mu_1 = \mu_2 = \mu_3$ contra la hipótesis alternativa de que el promedio de intervalo de atención es diferente para al menos uno de los tres tratamientos, se usa el estadístico F de análisis de varianza, calculado como

$$F = \frac{\text{MST}}{\text{MSE}} = \frac{29.2667}{5.9333} = 4.93$$

y se muestra en la columna marcada "F" de la figura 6.3a y "*F*" de la figura 6.3b. No es de sorprender saber que el valor en la columna marcada "P" en la figura 6.3a y "*valor P*" en la figura 6.3b sea el valor p exacto para esta prueba estadística.

La prueba estadística MST/MSE calculada líneas antes, tiene una distribución F con $gl_1 = 2$ y $gl_2 = 12$ grados de libertad. Con el uso del método del valor crítico con $\alpha = .05$, se puede rechazar H_0 si $F > F_{.05} = 3.89$ de la tabla 6 del Apéndice (véase la figura 6.6). Como el valor observado, $F = 4.93$, excede del valor crítico, se rechaza H_0. Hay suficiente evidencia para indicar que al menos uno de los tres intervalos de atención promedio es diferente de al menos uno de los otros.

FIGURA 6.6
Región de rechazo para el ejemplo 6.6



TIP MI CONSEJO

Las salidas impresas de computadora dan el valor *p* exacto; use el valor *p* para tomar su decisión.

Se podría haber llegado a esta misma conclusión usando el valor *p* exacto, $P = .027$, dado en la figura 6.3. Como el valor *p* es menor que $\alpha = .05$, los resultados son estadísticamente significativos al nivel de 5%. Todavía se concluye que al menos uno de los tres promedios de intervalos de atención es diferente de al menos uno de los otros.

NECESITO SABER...

Cómo determinar si mis cálculos son precisos

Las siguientes sugerencias se aplican a todos los análisis de varianza de este capítulo:

1. Cuando calcule sumas de cuadrados, asegúrese de llevar al menos seis cifras significativas antes de hacer restas.
2. Recuerde, las sumas de cuadrados nunca pueden ser negativas. Si obtiene una suma de cuadrados negativa, ha cometido un error aritmético.
3. Siempre verifique su tabla de análisis de varianza para comprobar que los grados de libertad ascienden al total de grados de libertad $(n - 1)$ y que las sumas de cuadrados ascienden al TSS.

Estadística de prueba, regla de decisión

Numerosos experimentos son exploratorios por naturaleza. No hay nociones preconcebidas de los resultados y no se ha decidido (antes de realizar el experimento) de hacer comparaciones específicas de tratamiento. En lugar de esto, se desea clasificar las medias de tratamiento, determinar cuáles medias difieren e identificar conjuntos de medias para las cuales no hay evidencia de diferencia.

Una opción podría ser ordenar las medias muestrales de menor a mayor y a continuación efectuar pruebas *t* para medias adyacentes en el ordenamiento. Si dos medias difieren en más de

$$t_{\alpha/2}\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

se concluye que el par de medias poblacionales difiere. El problema con este procedimiento es que la probabilidad de hacer un error tipo I, es decir, concluir que dos medias difieren cuando en realidad son iguales, es α para cada prueba. Si se compara un gran número de pares de medias, la probabilidad de detectar al menos una diferencia en medias, cuando en realidad ninguna existe, es bastante grande.

Una forma sencilla de evitar el alto riesgo de declarar diferencias cuando no existen es usar el **rango estudentizado**, que es la diferencia entre la más pequeña y la más grande en un conjunto de k medias muestrales, como la medida para determinar si hay una diferencia en un par de medias poblacionales. Este método, a veces denominado **método de Tukey para comparaciones pareadas**, hace que sea igual a α la probabilidad de declarar que existe una diferencia entre al menos un par en un conjunto de k medias de tratamiento, cuando no existe diferencia.

El método de Tukey para hacer comparaciones pareadas está basado en el análisis usual de suposiciones de varianza. **Además, supone que las medias muestrales son independientes y están basadas en muestras de igual tamaño.** La medida que determina si existe una diferencia entre un par de medias de tratamiento es la cantidad ω (omega minúscula), que se presenta a continuación.

MEDIDA PARA HACER COMPARACIONES PAREADAS

$$\omega = q_\alpha(k, gl)\left(\frac{s}{\sqrt{n_t}}\right)$$

donde

k = Número de tratamientos
s^2 = MSE = Estimador de la varianza común σ^2 y $s = \sqrt{s^2}$
gl = Número de grados de libertad para s^2
n_t = Tamaño muestral común, es decir, el número de observaciones en cada una de las k medias de tratamiento
$q_\alpha(k,gl)$ = Valor tabulado de las tablas 11a) y 11b) del Apéndice, para $\alpha = .05$ y .01, respectivamente, y para varias combinaciones de k y gl

Regla: Se juzga que dos medias poblacionales difieren si las medias muestrales correspondientes difieren en ω o más.

Las tablas 11a y 11b del Apéndice contienen los valores de $q_\alpha(k, gl)$ para $\alpha = .05$ y .01, respectivamente. Para ilustrar el uso de las tablas, consulte la parte de la tabla 11a) reproducida en la tabla 6.2. Suponga que se desea hacer comparaciones por pares de $k = 5$ medias con $\alpha = .05$ para un análisis de varianza, donde s^2 posee 9 gl. El valor tabulado para $k = 5$, $gl = 9$ y $\alpha = .05$, sombreados en la tabla 6.2, es $q_{.05}(5, 9) = 4.76$.

TABLA 6.2 Reproducción parcial de la tabla 11a del Apéndice; puntos de 5% superiores

gl	2	3	4	5	6	7	8	9	10	11	12
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61

EJEMPLO 6.7

Consulte el ejemplo 6.4, en el que se comparó el promedio de intervalos de atención para estudiantes a los que se dieron diferentes tratamientos de "comida" en la mañana: sin desayuno, un desayuno ligero o un desayuno completo. La prueba F de ANOVA del ejemplo 6.5 indicó una diferencia significativa en las medias poblacionales. Use el método de Tukey para comparaciones pareadas, para determinar cuál de las tres medias poblacionales difiere de las otras.

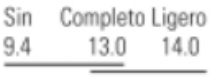
Solución Para este ejemplo, hay $k = 3$ medias de tratamiento, con $s = \sqrt{\text{MSE}} = 2.436$. El método de Tukey se puede usar, con cada una de las tres muestras conteniendo $n_t = 5$ mediciones y $(n - k) = 12$ grados de libertad. Consulte la tabla 11 en el Apéndice para hallar $q_{.05}(k, gl) = q_{.05}(3, 12) = 3.77$ y calcule la "regla" como

$$\omega = q_{.05}(3, 12)\left(\frac{s}{\sqrt{n_t}}\right) = 3.77\left(\frac{2.436}{\sqrt{5}}\right) = 4.11$$

Las tres medias de tratamiento están dispuestas en orden de la más pequeña, 9.4, a la máxima, 14.0, en la figura 6.7. El siguiente paso es comprobar la diferencia entre cada par de medias. La única diferencia que excede de $\omega = 4.11$ es la diferencia entre no desayuno y un desayuno ligero. Estos dos tratamientos se declaran, por tanto, significativamente diferentes. No se puede declarar una diferencia entre los otros dos pares de tratamientos. Para indicar este hecho

visualmente, la figura 6.7 muestra una línea bajo los pares de medias que no son significativamente diferentes.

FIGURA 6.7
Medias clasificadas para el ejemplo 6.7



Estos resultados pueden parecer confusos, pero por lo general ayuda considerar la clasificación de las medias e interpretar diferencias no significativas como nuestra incapacidad para clasificar de manera distintiva las medias subrayadas por la misma línea. Para este ejemplo, el desayuno ligero definitivamente clasificó más alto que sin desayuno, pero el desayuno completo no pudo ser clasificado más alto que sin desayuno, o más abajo que el desayuno ligero. La probabilidad de que cometimos al menos un error entre las tres comparaciones es a lo sumo $\alpha = .05$.

TIP MI CONSEJO

Si cero no está en el intervalo, hay evidencia de una diferencia entre los dos métodos

Muchos programas de computadora dan una opción para realizar **comparaciones pareadas**, incluyendo el método de Tukey, aunque *MS Excel* no la tiene. La salida *MINITAB* de la figura 6.8 muestra su forma de la prueba de Tukey, que difiere ligeramente del método que hemos presentado. Los tres intervalos que se ven en la salida impresa marcada "Inferior" y "Superior" representan la diferencia en las dos medias muestrales más o menos la regla ω. Si el intervalo contiene el valor 0, las dos medias se juzgan como no significativamente diferentes. Se puede ver que las únicas medias 1 y 2 (sin desayuno contra desayuno ligero) muestran una diferencia significativa.

FIGURA 6.8
Salida *MINITAB* para el ejemplo 6.7

```
Tukey's 95% Simultaneous Confidence Intervals
All Pairwise Comparisons among Levels of Meal
Individual confidence level = 97.94%

Meal = 1 subtracted from:

Meal  Lower  Center  Upper  -----+---------+---------+---------+----
2     0.493   4.600  8.707                (-----------*-----------)
3    -0.507   3.600  7.707              (----------*-----------)
                           -----+---------+---------+---------+----
                             -3.5       0.0       3.5       7.0
Meal = 2 subtracted from:

Meal  Lower  Center  Upper   -----+---------+---------+---------+----
3    -5.107  -1.000  3.107   (-----------*-----------)
                             -----+---------+---------+---------+----
                               -3.5       0.0       3.5       7.0
```

Cuando usted estudie dos diseños experimentales más en las siguientes secciones de este capítulo recuerde que, una vez que haya encontrado que un factor es significativo, debe usar el método de Tukey u otro método de comparaciones pareadas para averiguar exactamente dónde están las diferencias.

6.3 ANÁLISIS DE VARIANZA DE DOS VÍAS

El *diseño completamente aleatorizado* que se estudió en la sección 6.2 es una generalización del diseño de las *dos muestras independientes* presentado en el capítulo 4. Está pensado para usarse cuando las unidades experimentales sean bastante similares u *homogéneas* en su conformación y cuando haya sólo un factor, el *tratamiento*, que pueda influir en la respuesta.

Cualquier otra variación en la respuesta se debe a variación aleatoria o a *error experimental*. A veces es claro para el investigador que las unidades experimentales *no sean homogéneas*. Personas o animales con carácter experimental, campos agrícolas, días de la semana y otras unidades experimentales a veces agregan su propia variabilidad a la respuesta. Aun cuando el investigador no esté realmente interesado en esta fuente de variación, sino que más bien en algún *tratamiento* que elija aplicar, puede aumentar la información al aislar esta fuente de

variación usando el **diseño de bloque aleatorizado**, que es una extensión directa de los *pares acoplados* o *diseño de diferencia pareada* de la del capítulo 4.

Pruebas de hipótesis

TIP MI CONSEJO

b = bloques
k = tratamientos
$n = bk$

En un diseño de bloque aleatorizado, el experimentador está interesado en comparar k medias de tratamiento. El diseño utiliza *bloques* de k unidades experimentales que son relativamente similares, u *homogéneos*, con una unidad dentro de cada bloque asignada *al azar* a cada tratamiento. Si el diseño de bloque aleatorizado contiene k tratamientos dentro de cada uno de los b bloques, entonces el número total de observaciones en el experimento es $n = bk$.

Un supervisor de producción desea comparar los tiempos medios para que operadores de línea de producción ensamblen un artículo usando uno de tres métodos: A, B o C. Esperando variación en tiempos de ensamble de un operador a otro, el supervisor emplea un diseño de bloque aleatorizado para comparar los tres métodos. Cinco operadores de línea de producción se seleccionan para servir como bloques y a cada uno se le asigna ensamblar el artículo tres veces, una vez por cada uno de los tres métodos. Como la secuencia en la que el operador usa los tres métodos puede ser importante (fatiga o mayor destreza pueden ser factores que afecten la respuesta), cada operador debe ser asignado a una secuencia aleatoria de los tres métodos. Por ejemplo, el operador 1 podría ser asignado a efectuar primero el método C, seguido por A y B. El operador 2 podría realizar primero el método A, seguido del C y B, etcétera.

Para comparar cuatro métodos diferentes de enseñanza, un grupo de estudiantes podría ser dividido en bloques de tamaño 4, de modo que los grupos se encuentren casi *acoplados* según su rendimiento académico. Para comparar los costos promedio de tres compañías de telefonía celular diferentes, los costos podrían compararse en cada uno de tres niveles de uso: bajo, medio y alto. Para comparar el promedio de producción de tres especies de árboles frutales, cuando se espera una variación en producción debido al campo en el que se planten los árboles, una investigadora usa cinco campos. Ella divide cada campo en tres *lotes* en los que están plantadas las tres especies de árboles frutales.

El acoplamiento o *bloqueo* puede tener lugar en muchas formas diferentes. Las comparaciones de tratamientos se hacen a veces dentro de bloques de tiempo, dentro de bloques de personas o dentro de ambientes externos similares. El propósito de bloquear es remover o aislar la variabilidad *de un bloque a otro* que pudiera de otro modo ocultar el efecto de los tratamientos.

Variaciones considerando grupos y bloques

División de la variación total en el experimento

Sea x_{ij} la respuesta cuando el i–ésimo tratamiento ($i = 1, 2, \ldots, k$) se aplica en el j–ésimo bloque ($j = 1, 2, \ldots, b$). La variación total en las $n = bk$ observaciones es

$$\text{TSS} = \Sigma(x_{ij} - \bar{x})^2 = \Sigma x_{ij}^2 - \frac{(\Sigma x_{ij})^2}{n}$$

Esto se divide en *tres* (en lugar de dos) partes de modo que

$$\text{TSS} = \text{SSB} + \text{SST} + \text{SSE}$$

donde

- SSB (suma de cuadrados para bloques) mide la variación entre las medias de bloque.
- SST (suma de cuadrados para tratamientos) mide la variación entre las medias de tratamiento.
- SSE (suma de cuadrados para error) mide la variación de las diferencias entre las observaciones de tratamiento *dentro* de bloques, que mide el error experimental.

Las fórmulas de cálculo para las cuatro sumas de cuadrados son semejantes en forma a las empleadas para el diseño completamente aleatorizado de la sección 6.2. Aun cuando se puede

simplificar el trabajo con el uso de un programa computarizado para calcular estas sumas de cuadrados, las fórmulas se dan a continuación.

CÁLCULO DE LAS SUMAS DE CUADRADOS PARA UN DISEÑO DE BLOQUE ALEATORIZADO, *k* TRATAMIENTOS EN *b* BLOQUES

TIP MI CONSEJO

TSS = SST + SSB + SSE

$$\text{CM} = \frac{G^2}{n}$$

donde

$$G = \Sigma x_{ij} = \text{Total de todas las } n = bk \text{ observaciones}$$

$$\text{TSS} = \Sigma x_{ij}^2 - \text{CM}$$

$$= (\text{suma de todos los valores de } x) - \text{CM}$$

$$\text{SST} = \Sigma \frac{T_i^2}{b} - \text{CM}$$

$$\text{SSB} = \Sigma \frac{B_j^2}{k} - \text{CM}$$

$$\text{SSE} = \text{TSS} - \text{SST} - \text{SSB}$$

con

T_i = Total de todas las observaciones que reciben tratamiento i, $i = 1, 2, \ldots, k$

B_j = Total de todas las observaciones en bloque j, $j = 1, 2, \ldots, b$

Cada una de las tres **fuentes de variación**, cuando están divididas por los **grados de libertad** apropiados, da una estimación de la variación en el experimento. Como TSS contiene $n = bk$ observaciones al cuadrado, sus grados de libertad son $gl = (n - 1)$. Del mismo modo, SST contiene k totales al cuadrado y sus grados de libertad son $gl = (k - 1)$, en tanto que SSB contiene b totales al cuadrado y $(b - 1)$ grados de libertad. Finalmente, como los grados de libertad son aditivos, los restantes grados de libertad asociados con SSE se puede demostrar algebraicamente que son $gl = (b - 1)(k - 1)$.

Estas tres fuentes de variación y sus respectivos grados de libertad se combinan para formar las **medias cuadráticas** como MS = SS/gl, y la variación total en el experimento se exhibe entonces en una **tabla de análisis de varianza** (o **ANOVA**) como se muestra a continuación:

TIP MI CONSEJO

Los grados de libertad son aditivos

ANOVA PARA UN DISEÑO DE BLOQUE ALEATORIZADO, *k* TRATAMIENTOS Y *b* BLOQUES

Fuente	*gl*	SS	MS	*F*
Tratamientos	$k - 1$	SST	MST = SST/$(k - 1)$	MST/MSE
Bloques	$b - 1$	SSB	MSB = SSB/$(b - 1)$	MSB/MSE
Error	$(b - 1)(k - 1)$	SSE	MSE = SSE/$(b - 1)(k - 1)$	
Total	$n - 1 = bk - 1$			

EJEMPLO 6.8

La industria de la telefonía celular está inmersa en una feroz batalla por los clientes, con cada compañía ideando su propio y complejo plan de precios para atraer clientes. Como el costo de un minuto por teléfono celular varía en forma drástica dependiendo del número de minutos por mes usados por el cliente, un grupo de vigilancia integrado por consumidores decidió comparar el promedio de costos de cuatro compañías de telefonía celular usando tres diferentes niveles de uso como bloques. Los costos mensuales (en dólares), calculados por las compañías de telefonía celular para usuarios de tiempo pico en niveles de uso bajo (20 minutos por mes), medio (150 minutos por mes) y alto (1 000 minutos por mes), se dan en la tabla 6.3. Construya la tabla de análisis de varianza para este experimento.

TABLA 6.3 **Costos mensuales de teléfono**

	Compañía				
Nivel de uso	A	B	C	D	Totales
Bajo	27	24	31	23	$B_1 = 105$
Medio	68	76	65	67	$B_2 = 276$
Alto	308	326	312	300	$B_3 = 1246$
Totales	$T_1 = 403$	$T_2 = 426$	$T_3 = 408$	$T_4 = 390$	$G = 1627$

TIP MI CONSEJO

Los bloques contienen unidades experimentales que son *relativamente las mismas*

Solución El experimento está diseñado como *diseño de bloque aleatorizado* con $b = 3$ niveles de uso (bloques) y $k = 4$ compañías (tratamientos), de modo que hay $n = bk = 12$ observaciones y $G = 1\,627$. Entonces

$$\text{CM} = \frac{G^2}{n} = \frac{1627^2}{12} = 220\,594.0833$$

$$\text{TSS} = (27^2 + 24^2 + \cdots + 300^2) - \text{CM} = 189\,798.9167$$

$$\text{SST} = \frac{403^2 + \cdots + 390^2}{3} - \text{CM} = 222.25$$

$$\text{SSB} = \frac{105^2 + 276^2 + 1246^2}{4} - \text{CM} = 189\,335.1667$$

y por sustracción,

$$\text{SSE} = \text{TSS} - \text{SST} - \text{SSB} = 241.5$$

Estas cuatro fuentes de variación, sus grados de libertad, sumas de cuadrados y medias cuadráticas se muestran en el área sombreada de la tabla de análisis de varianza, generadas por *MINITAB* y *MS Excel* y se dan en las figuras 6.9a y 6.9b. Encontrará instrucciones para generar esta salida impresa en la sección 6.4.

FIGURA 6.9a

Salida *MINITAB* para el ejercicio 6.8

ANOVA de dos vías: dólares contra uso, compañía

Source	DF	SS	MS	F	P
Usage	2	189335	94667.6	2351.99	0.000
Company	3	222	74.1	1.84	0.240
Error	6	242	40.3		
Total	11	189799			

S = 6.344 R-Sq = 99.87% R-Sq(adj) = 99.77%

FIGURA 6.9b

Salida de *MS Excel* para el ejemplo 6.8

ANOVA: dos factores sin réplica

RESUMEN	*Conteo*	*Suma*	*Promedio*	*Varianza*
Bajo	4	105	26.25	12.917
Medio	4	276	69	23.333
Alto	4	1246	311.5	118.333
A	3	403	134.333	23040.333
B	3	426	142	26068
C	3	408	136	23521
D	3	390	130	22159

ANOVA

Fuente de variación	*SS*	*gl*	*MS*	*F*	*Valor P*	*F crit.*
Uso	189335.167	2	94667.583	2351.990	0.000	5.143
Compañía	222.25	3	74.083	1.841	0.240	4.757
Error	241.5	6	40.25			
Total	189798.917	11				

Observe que ambas tablas ANOVA muestran dos diferentes estadísticos F y valores de p. No debe sorprender el saber que estos estadísticos se usan para probar hipótesis respecto a la igualdad de medias *de tratamiento* y *de bloque*.

Prueba de la igualdad de las medias de tratamiento y de bloque

Las *medias cuadráticas* de la tabla de análisis de varianza se pueden usar para probar las hipótesis nulas

H_0: No hay diferencia entre las k medias de tratamiento

o bien

H_0: No hay diferencia entre las b medias de bloque

contra la hipótesis alternativa

H_a: Al menos una de las medias es diferente de al menos otra

con el uso de un argumento teórico similar al empleado para el diseño completamente aleatorizado.

- Recuerde que σ^2 es la varianza común para las observaciones en todas las bk combinaciones de bloque–tratamiento. La cantidad

$$\text{MSE} = \frac{\text{SSE}}{(b-1)(k-1)}$$

es una estimación insesgada de σ^2, sea H_0 verdadera o no.

- Las dos medias cuadráticas, MST y MSB, estiman σ^2 sólo si H_0 es verdadera y tienden a ser inusualmente *grandes* si H_0 es falsa y ya sea que las medias de tratamiento o de bloque sean diferentes.

Estadísticos de prueba, regla de decisión

Los estadísticos de prueba

$$F = \frac{\text{MST}}{\text{MSE}} \quad \text{y} \quad F = \frac{\text{MSB}}{\text{MSE}}$$

se usan para probar la igualdad de medias de tratamiento y de bloque, respectivamente. Ambos estadísticos tienden a ser más grandes de lo normal si H_0 es falsa. En consecuencia, se puede rechazar H_0 para valores grandes de F, usando valores críticos *de cola derecha* de la distribución F con los grados de libertad apropiados (véase la tabla 6 del Apéndice) o valores p generados por computadora para sacar conclusiones estadísticas acerca de la igualdad de las medias poblacionales.

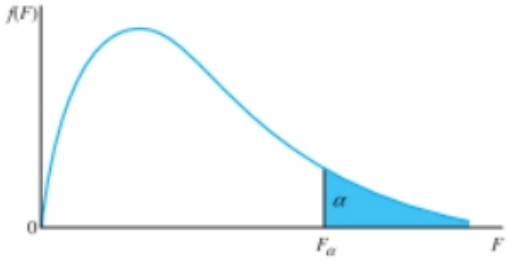
PRUEBAS PARA UN DISEÑO DE BLOQUE ALEATORIZADO

Para comparar medias de tratamiento:

1. Hipótesis nula: H_0: las medias de tratamiento son iguales
2. Hipótesis alternativa: H_a: al menos dos de las medias de tratamiento difieren
3. Prueba estadística: $F = \text{MST/MSE}$, donde F está basada en $gl_1 = (k-1)$ y $gl_2 = (b-1)(k-1)$
4. Región de rechazo: rechazar si $F > F_\alpha$, donde F_α se encuentra en la cola superior de la distribución F (véase la figura), o cuando el valor $p < \alpha$

Para comparar medias de bloque:

1. Hipótesis nula: H_0: las medias de bloque son iguales
2. Hipótesis alternativa: H_a: al menos dos de las medias de bloque difieren
3. Prueba estadística: $F = \text{MSB/MSE}$, donde F está basada en $gl_1 = (b - 1)$ y $gl_2 = (b - 1)(k - 1)$
4. Región de rechazo: rechazar si $F > F_\alpha$, donde F_α se encuentra en la cola superior de la distribución F (véase la figura), o cuando el valor $p < \alpha$



EJEMPLO 6.9 ¿Los datos del ejemplo 6.8 dan suficiente evidencia para indicar una diferencia en el promedio de costo mensual de teléfono celular, dependiendo de la compañía que use el cliente?

Solución Las compañías de telefonía celular representan los *tratamientos* en este diseño de bloque aleatorizado y las diferencias en el promedio de sus costos mensuales son de primordial importancia para el investigador. Para probar

H_0: No hay diferencia en el promedio de costo entre compañías

contra la alternativa de que el promedio de costo es diferente para al menos una de las cuatro compañías, se usa el análisis de varianza del estadístico F, calculado como

$$F = \frac{\text{MST}}{\text{MSE}} = \frac{74.1}{40.3} = 1.84$$

y se muestra en la columna marcada "F" y el renglón marcado "Compañía" en las figuras 6.9a y 6.9b. El valor p exacto para esta prueba estadística también se da en las figuras 6.9a y 6.9b como .240, que es demasiado grande para permitir el rechazo de H_0. Los resultados no muestran una diferencia significativa en las medias de tratamiento. Esto es, hay insuficiente evidencia para indicar una diferencia en el promedio de costos mensuales para las cuatro compañías.

El investigador del ejemplo 6.9 estaba bastante seguro de que al usar un *diseño de bloque aleatorizado* habría una diferencia significativa en las medias de bloque, es decir, una diferencia significativa en el promedio de costos mensuales dependiendo del nivel de uso. Esta sospecha está justificada al ver la prueba de igualdad de medias de bloque. Observe que la prueba estadística observada es $F = 2\,351.99$ con valor $p = .000$, mostrando una diferencia altamente significativa, como se esperaba, en las medias de bloque.

Identificación de diferencias en las medias de tratamiento y de bloque

Una vez efectuada la prueba F general para igualdad de las medias de tratamiento o de bloque, ¿qué más se puede hacer para identificar la naturaleza de cualesquiera diferencias que se hayan encontrado? Al igual que en la sección 6.2, se puede usar el método de Tukey de comparaciones pareadas para determinar qué pares de medias de tratamiento o de bloque son significativamente diferentes uno de otro. No obstante, si la prueba F no indica una diferencia significativa en las medias, no hay razón para usar el procedimiento de Tukey. Si el investigador tiene especial interés en un *par* particular de medias de tratamiento o de bloque, puede estimar la diferencia

usando un intervalo de confianza $(1 - \alpha)100\%$.† Las fórmulas para estos procedimientos, que se muestran a continuación, siguen un patrón similar a las fórmulas para el diseño completamente aleatorizado. Recuerde que MSE siempre da un estimador insesgado de σ^2 y utiliza información proveniente de todo el conjunto de mediciones. En consecuencia, es el mejor estimador existente de σ^2, cualquiera que sea el procedimiento de prueba o estimación que se use. Utilice de nuevo

$$s^2 = \text{MSE} \qquad \text{con } gl = (b - 1)(k - 1)$$

para estimar σ^2 al comparar las medias de tratamiento y de bloque.

TIP MI CONSEJO

Los grados de libertad para la prueba de Tukey y para intervalos de confianza son los ***df de error***

COMPARACIÓN DE MEDIAS DE TRATAMIENTO Y DE BLOQUE

Medidor de Tukey para comparar medias de bloque:

$$\omega = q_\alpha(b, gl)\left(\frac{s}{\sqrt{k}}\right)$$

Medidor de Tukey para comparar medias de tratamiento:

$$\omega = q_\alpha(k, gl)\left(\frac{s}{\sqrt{b}}\right)$$

Intervalo de confianza de $(1 - \alpha)100\%$ para la diferencia en dos medias de bloque:

$$(\overline{B}_i - \overline{B}_j) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{k} + \frac{1}{k}\right)}$$

donde $\overline{B}_i$ es el promedio de todas las observaciones en el bloque i

Intervalo de confianza de $(1 - \alpha)100\%$ para la diferencia en dos medias de tratamiento:

$$(\overline{T}_i - \overline{T}_j) \pm t_{\alpha/2}\sqrt{s^2\left(\frac{1}{b} + \frac{1}{b}\right)}$$

donde $\overline{T}_i$ es el promedio de todas las observaciones en el tratamiento i.

Nota: Los valores $q_\alpha(*, gl)$ de la tabla 11 en el Apéndice, $t_{\alpha/2}$ de la tabla 4 en el Apéndice y $s^2 = \text{MSE}$ dependen todos ellos de $gl = (b - 1)(k - 1)$ grados de libertad.

EJEMPLO 6.10

Identifique la naturaleza de cualesquiera diferencias encontradas en el promedio de costos mensuales de teléfono celular del ejemplo 6.8.

Solución Como la prueba F no presentó diferencias significativas en los costos promedio para las cuatro compañías, no hay razón para usar el método de Tukey de comparaciones pareadas. Supongamos, sin embargo, que usted es un ejecutivo de la compañía B y su principal competidor es la compañía C. ¿Puede decir que hay una diferencia significativa en los dos costos promedio? Usando un intervalo de confianza de 95%, puede calcular

TIP MI CONSEJO

No se puede formar un intervalo de confianza o probar una hipótesis acerca de una media de tratamiento individual en un diseño de bloque aleatorizado

$$(\overline{T}_2 - \overline{T}_3) \pm t_{.025}\sqrt{\text{MSE}\left(\frac{2}{b}\right)}$$

$$\left(\frac{426}{3} - \frac{408}{3}\right) \pm 2.447\sqrt{40.3\left(\frac{2}{3}\right)}$$

$$6 \pm 12.68$$

† No se puede construir un intervalo de confianza para una media individual, a menos que los bloques hayan sido seleccionados al azar de entre la población de todos los bloques. El procedimiento para construir intervalos para medias individuales está fuera del propósito de este libro.

de modo que la diferencia entre los dos costos promedio se estima entre –$6.68 y $18.68. Como 0 está contenido en el intervalo, no tiene evidencia para indicar una diferencia significativa en sus costos promedio. ¡Qué pena!

UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

6.4

Procedimientos de análisis de varianza: *Microsoft Excel*

Los procedimientos estadísticos para realizar el análisis de varianza para los tres diseños experimentales que se expusieron en este capítulo pueden encontrarse usando el comando de *Microsoft Excel* **Data ▶ Data Analysis**. Verá opciones para **Single Factor**, **Two–Factor Without Replication** y **Two–Factor With Replication** que generarán cuadros de diálogo usados para los diseños completamente aleatorizado, de bloque aleatorizado y factorial, respectivamente.

EJEMPLO 6.11

(Diseño completamente aleatorizado) Consulte el estudio del desayuno en el ejemplo 6.4, en el cual se estudió el efecto de la nutrición en el intervalo de atención.

Sin desayuno	Desayuno ligero	Desayuno completo
8	14	10
7	16	12
9	12	16
13	17	15
10	11	12

Introduzca los datos en las columnas A, B y C de una hoja de cálculo de *Excel* con una muestra por columna.

1. Use **Data ▶ Data Analysis ▶ Anova: Single Factor** para generar el cuadro de diálogo de la figura 6.10a. Resalte o teclee el **Input Range** (los datos en las primeras tres columnas) en el primer cuadro. En la sección marcada "Grouped by" elija el botón de radio para **Columns** y seleccione "Labels" si es necesario.
2. En *Excel*, el nivel de significancia por omisión es $\alpha = .05$. Cambie este nivel de significancia si es necesario. Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK**. La salida aparecerá en la ubicación de celda seleccionada, y deberá ajustarse usando **Format ▶ AutoFit Column Width** en la pestaña **Home** en el grupo **Cells** mientras todavía está resaltado. Puede disminuir la precisión decimal si lo desea, usando en la pestaña **Home** en el grupo **Number** (véase la figura 6.10b).
3. El valor observado del estadístico de prueba $F = 4.93$ se encuentra en la fila etiquetada "*Between Groups*" seguido por el "*P–value*" y el valor crítico marcando la región de rechazo para una prueba de una cola con $\alpha = .05$. Para este ejemplo, el valor $p = .0273$ indica que hay una diferencia significativa en los intervalos de atención promedio dependiendo del tipo de desayuno.

FIGURA 6.10

(a)

(b)

E	F	G	H	I	J	K
Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
None	5	47	9.4	5.3		
Light	5	70	14	6.5		
Full	5	65	13	6		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	58.533	2	29.2667	4.93258	0.0273	3.89
Within Groups	71.2	12	5.93333			
Total	129.73	14				

EJERCICIOS

TÉCNICAS BÁSICAS

6.1 Suponga que se desea comparar las medias de seis poblaciones basadas en muestras aleatorias independientes, cada una de las cuales contiene 10 observaciones. Inserte, en una tabla ANOVA, las fuentes de variación y sus respectivos grados de libertad.

DATOS EX1107

6.2 Estos datos son observaciones recolectadas usando un diseño completamente aleatorizado:

a. Calcule CM y TSS.

b. Calcule SST y MST.

c. Calcule SSE y MSE.

d. Construya una tabla ANOVA para los datos.

e. Exprese las hipótesis nula y alternativa para un análisis de prueba F de varianza.

f. Use el método del valor p para determinar si hay una diferencia en las tres medias poblacionales.

6.3 Si el tamaño muestral para cada tratamiento es n_t y si s^2 está basada en 12 gl, encuentre ω en estos casos:

a. $\alpha = .05, k = 4, n_t = 5$

b. $\alpha = .01, k = 6, n_t = 8$

DATOS EX1133

6.4 Los datos que siguen son observaciones recolectadas de un experimento que comparó cuatro tratamientos, A, B, C y D, dentro de cada uno de tres bloques, usando un diseño de bloque aleatorizado.

	Tratamiento				
Bloque	A	B	C	D	Total
1	6	10	8	9	33
2	4	9	5	7	25
3	12	15	14	14	55
Total	22	34	27	30	113

a. ¿Los datos presentan suficiente evidencia para indicar diferencias entre las medias de tratamiento? Pruebe usando $\alpha = .05$.

b. ¿Los datos presentan suficiente evidencia para indicar diferencias entre las medias de bloque? Pruebe usando $\alpha = .05$.

c. Clasifique las cuatro medias de tratamiento usando el método de Tukey de comparaciones pareadas con $\alpha = .01$.

d. Encuentre un intervalo de confianza de 95% para la diferencia en medias para los tratamientos A y B.

e. ¿Le parece que el uso de un diseño de bloque aleatorizado para este experimento estaba justificado? Explique.

APLICACIONES

DATOS EX1110

6.5 Reducir hostilidad Un psicólogo clínico deseaba comparar tres métodos para reducir niveles de hostilidad en estudiantes universitarios, con el uso de cierto examen psicológico (HLT). Se tomaron las calificaciones altas de este examen como indicio de gran hostilidad. Once estudiantes que obtuvieron calificaciones altas y casi iguales se emplearon en el experimento. Cinco fueron seleccionados al azar de entre los 11 estudiantes y tratados con el método A, tres se tomaron al azar de los seis estudiantes restantes y se trataron con el método B, y los otros tres estudiantes con el método C. Todos los tratamientos continuaron durante todo un semestre, cuando el examen HLT se aplicó de nuevo. Los resultados se muestran en la tabla.

Método	Calificaciones en el examen HLT				
A	73	83	76	68	80
B	54	74	71		
C	79	95	87		

a. Realice un análisis de varianza para este experimento.

b. ¿Los datos dan suficiente evidencia para indicar una diferencia en respuesta media de estudiantes a los tres métodos después del tratamiento?

DATOS EX1115

6.6 Calcio El contenido de calcio de una sustancia mineral pulverizada fue analizado cinco veces por cada uno de tres métodos, con desviaciones estándar similares:

Método	Porcentaje de calcio				
1	.0279	.0276	.0270	.0275	.0281
2	.0268	.0274	.0267	.0263	.0267
3	.0280	.0279	.0282	.0278	.0283

Use una prueba adecuada para comparar los tres métodos de medición. Comente sobre la validez de cualesquiera suposiciones que sea necesario hacer.

DATOS EX1117

6.7 El costo de la madera Un constructor de viviendas a nivel nacional desea comparar los precios por 1 000 pies de madera de armazones de abeto Douglas de calidad estándar o mejor. Al azar selecciona cinco proveedores en cada uno de los cuatro estados donde el constructor planea iniciar la construcción. Los precios se dan en la tabla siguiente.

Estado			
1 ($)	2 ($)	3 ($)	4 ($)
261	236	250	265
255	240	245	270
258	225	255	258
267	233	248	275
270	240	260	275

a. ¿Qué tipo de diseño experimental se ha empleado?

b. Construya la tabla del análisis de varianza para estos datos.

c. ¿Los datos dan suficiente evidencia para indicar que el precio promedio por 1 000 pies de madera de abeto Douglas difiere entre los cuatro estados? Pruebe usando $\alpha = .05$.

DATOS EX1125

6.8 Tolerancia a la glucosa Los médicos dependen de resultados de exámenes de laboratorio cuando manejan problemas médicos como diabetes o epilepsia. En un examen de uniformidad para tolerancia a la glucosa, a tres laboratorios diferentes se les enviaron $n_t = 5$ muestras de sangre idénticas de una persona que había bebido 50 miligramos (mg) de glucosa

disuelta en agua. Los resultados de laboratorio (en mg/dl) son los siguientes:

Lab 1	Lab 2	Lab 3
120.1	98.3	103.0
110.7	112.1	108.5
108.9	107.7	101.1
104.2	107.9	110.0
100.4	99.2	105.4

a. ¿Los datos indican una diferencia en el promedio de lecturas para los tres laboratorios?

b. Use el método de Tukey para comparaciones pareadas para clasificar las tres medias de tratamiento. Use $\alpha = .05$.

DATOS EX1136

6.9 Rendimiento de gasolina Se realizó un estudio para comparar el rendimiento de gasolina en automóviles para tres fórmulas de gasolina. Se utilizaron cuatro automóviles, todos ellos de la misma marca y modelo, y cada fórmula se probó en cada uno de los autos. El uso de cada fórmula en el mismo auto tiene el efecto de eliminar (bloquear) variabilidad de un auto a otro. Los datos (en millas por galón) se dan a continuación.

	Automóvil			
Fórmula	1	2	3	4
A	25.7	27.0	27.3	26.1
B	27.2	28.1	27.9	27.7
C	26.1	27.5	26.8	27.8

a. ¿Los datos presentan suficiente evidencia para indicar diferencias en distancia media en millas por galón para las tres fórmulas de gasolina?

b. ¿Hay suficiente evidencia de una diferencia en distancia media en millas para los cuatro automóviles?

c. Suponga que *antes de ver los datos*, se ha decidido comparar la distancia media en millas por galón para las fórmulas A y B. Encuentre un intervalo de confianza de 90% para esta diferencia.

d. Use un método apropiado para identificar las diferencias por pares, si las hay, en el promedio de las distancias medias para las tres fórmulas.

DATOS EX1143

6.10 ¿Dónde comprar? ¿Usted compra en la tienda de abarrotes más cercana a su casa o busca la tienda que tenga los mejores precios? Se compararon los precios normales en cuatro tiendas de abarrotes diferentes para ocho artículos comprados el mismo día.

	Tiendas			
Artículos	Vons	Ralphs	Stater Bros	WinCo
Ensalada mixta, bolsa de 12 oz.	3.99	2.79	1.99	1.78
Salchichas de res ahumadas Hillshire Farm®, 14 oz.	4.29	4.29	3.99	2.50
Raisin Bran® de Kellogg's, 25.5 oz.	4.49	5.49	4.49	3.15
Queso crema Philadelphia® de Kraft®, 8 oz.	2.99	3.19	2.79	1.48
Aderezo Ranch de Kraft®, 16 oz.	3.19	3.49	3.49	1.48
Jugo de manzana Treetop®, 64 oz.	2.99	3.49	3.49	1.58
Jabón Dial®, Gold, 8–4 oz.	5.99	6.49	5.79	5.14
Crema de cacahuate Jif®, cremosa, 28 oz.	5.15	5.49	4.79	4.34

a. ¿Cuáles son los bloques y tratamientos en este experimento?

b. ¿Los datos dan evidencia para indicar que hay diferencias significativas en precios de una tienda a otra? Apoye su respuesta estadísticamente usando la salida impresa ANOVA que sigue.

c. ¿Hay diferencias significativas de un bloque a otro? ¿Fue eficaz el bloqueo?

ANOVA de dos vías: precio contra artículo, tienda

```
Fuente    GL        SS         MS        F      P
Artículo  7    40.2184    5.74548    29.99  0.000
Tienda    3    14.6695    4.88982    25.53  0.000
Error     21    4.0230    0.19157
Total     31   58.9108

S = 0.4377  R-Sq = 93.17%  R-Sq(adj) = 89.92%
```

6.11 ¿Dónde comprar?, continúa Consulte el ejercicio 6.10. La salida impresa que sigue da el promedio de costos de los artículos seleccionados para las $k = 4$ tiendas.

Tienda	Media
Vons	4.1350
Ralphs	4.3400
Stater	3.8525
WinCo	2.5975

a. ¿Cuál es el valor apropiado de $q_{.05}(k, gl)$ para probar diferencias entre tiendas?

b. ¿Cuál es el valor de $\omega = q_{.05}(k, gl)\sqrt{\frac{MSE}{b}}$?

c. Use la prueba de comparación por pares de Tukey entre tiendas usadas para determinar cuáles tiendas difieren significativamente en promedio de precios de los artículos seleccionados.

DATOS EX1156

6.12 Tiempos de reacción contra estímulos Veintisiete personas participaron en un experimento para comparar los efectos de cinco diferentes estímulos sobre el tiempo de reacción. El experimento se corrió usando un diseño completamente aleatorizado y, cualquiera que fuera el resultado del análisis de varianza, los experimentadores deseaban comparar los estímulos A y D. Los resultados del experimento se dan aquí. Use la salida impresa *MINITAB* para completar el ejercicio.

Estímulo	Tiempo de reacción (sec)							Total	Media
A	.8	.6	.6	.5				2.5	.625
B	.7	.8	.5	.5	.6	.9	.7	4.7	.671
C	1.2	1.0	.9	1.2	1.3	.8		6.4	1.067
D	1.0	.9	.9	1.1	.7			4.6	.920
E	.6	.4	.4	.7	.3			2.4	.480

Salida *MINITAB* para el ejercicio 6.12

Anova en una vía: tiempo contra estímulo

```
Source     DF       SS        MS        F      P
Stimulus    4   1.2118    0.3030    11.67  0.000
Error      22   0.5711    0.0260
Total      26   1.7830

S = 0.1611   R-Sq = 67.97%   R-Sq(adj) = 62.14%

                                Individual 95% CIs For Mean Based on
                                Pooled StDev
Level      N     Mean    StDev  ---------+---------+---------+---------+--
A          4   0.6250   0.1258        (------*------)
B          7   0.6714   0.1496          (----*----)
C          6   1.0667   0.1966                        (-----*-----)
D          5   0.9200   0.1483                    (-----*-----)
E          5   0.4800   0.1643  (-----*-----)
                                ---------+---------+---------+---------+--
Pooled StDev =   0.1611              0.50      0.75      1.00      1.25
```

a. Realice un análisis de varianza y pruebe por si hay diferencia en los tiempos medios de reacción debido a los cinco estímulos.

b. Compare los estímulos A y D para ver si hay una diferencia en tiempos medios de reacción.

6.13 Factorial de 4 × 2 Suponga que usted debe realizar un experimento factorial de dos factores, el factor A en cuatro niveles y el factor B en dos niveles, con r réplicas por tratamiento.

a. ¿Cuántos tratamientos intervienen en el experimento?

b. ¿Cuántas observaciones están involucradas?

c. Haga una lista de las fuentes de variación y sus grados de libertad respectivos.

DATOS EX1170

6.14 Salarios de profesores En un estudio sobre los salarios iniciales de los profesores auxiliares[1] cinco profesores varones y cinco mujeres que iniciaban en cada uno de dos tipos de instituciones que otorgan títulos de doctorado participaron en una encuesta y se registraron sus salarios iniciales. Los resultados de la encuesta en unidades de miles de dólares se dan en la tabla siguiente.

	Tipo	
Género	Público ($)	Sin fines de lucro ($)
Male	64.6	75.2
	63.8	75.7
	63.3	75.0
	64.1	75.1
	64.9	75.8
Female	59.7	71.6
	56.9	71.6
	58.4	71.2
	60.1	73.0
	59.8	69.5

a. ¿Qué tipo de diseño se utilizó para recolectar estos datos?

b. Use un análisis de varianza para probar si hay diferencias significativas en género y tipo de institución, y para probar si hay interacción significativa de género × tipo de institución.

c. Encuentre una estimación de intervalo de confianza de 95% para la diferencia en salarios iniciales para profesores auxiliares hombres y profesoras auxiliares mujeres. Interprete este intervalo en términos de una diferencia de género en salarios iniciales.

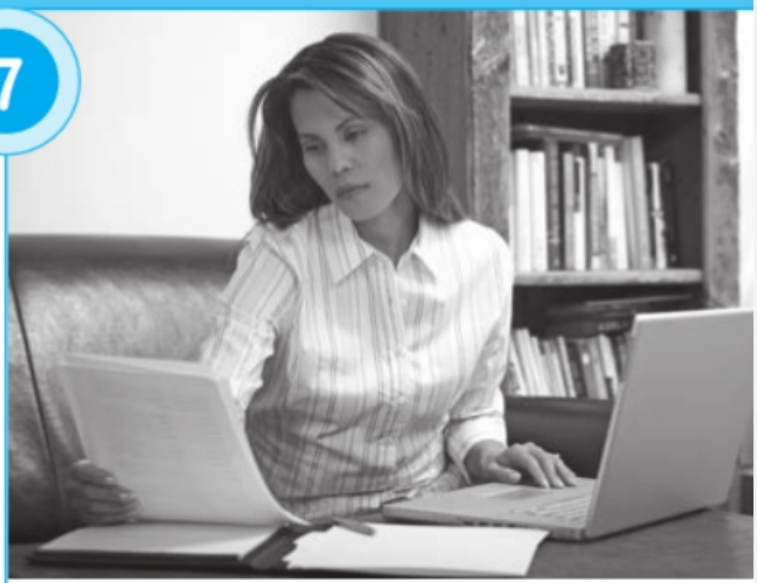
d. Elabore un resumen de los resultados de su análisis.

Fuente

1. Basado en datos de "Average Salary of Full Time Instructional Faculty on 9-Month Contracts in Degree-granting Institutions," *Digest of Educational Statistics*, http://nces.ed.gov/programs/digest/d09/tables/dt09_257.asp, 5 de julio de 2010.

7

Pruebas de chi cuadrada



© rubberball/Jupiter Images

OBJETIVOS GENERALES

Numerosos tipos de estudios y experimentos resultan en variables de respuesta cualitativas y no cuantitativas, de modo que las respuestas pueden ser clasificadas pero no cuantificadas. Los datos de estos experimentos están formados por el conteo o número de observaciones que caen en cada una de las categorías de respuesta incluidas en el experimento. En este capítulo, nos ocupamos de métodos para analizar datos categóricos.

ÍNDICE DEL CAPÍTULO

- Comparación de varias poblaciones multinomiales
- Otras aplicaciones
- Tablas de contingencia

NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

7.1 PRUEBAS DE CHI CUADRADA PARA LA DIFERENCIA ENTRE DOS PROPORCIONES (MUESTRAS INDEPENDIENTES)

En algunas situaciones, el investigador clasifica una unidad experimental de acuerdo con *dos variables cualitativas* para generar *datos bivariados*.

- Una pieza defectuosa de mueble se clasifica según el tipo de defecto y el turno de producción durante el que se hizo.
- Una profesora es clasificada por su rango profesional y el tipo de universidad (pública o privada) en la que trabaja.
- Un paciente es clasificado de acuerdo con el tipo de tratamiento preventivo contra la gripe que ha recibido y si ha contraído o no la gripe durante el invierno.

Cuando se registran dos *variables categóricas*, se pueden resumir los datos al contar el número observado de unidades que caen en cada una de las diversas intersecciones de niveles de categoría. Las cantidades resultantes se exhiben en un conjunto ordenado llamado **tabla de contingencia**.

EJEMPLO 7.1

Un total de $n = 309$ defectos en muebles fueron registrados y clasificados en cuatro tipos: A, B, C o D. Al mismo tiempo, cada pieza de mueble fue identificada por el turno de producción en el que se manufacturó. Estas cantidades se presentan en una tabla de contingencia en la tabla 7.1.

TABLA 7.1 Tabla de contingencia

	Turno			
Tipo de defectos	1	2	3	Total
A	15	26	33	74
B	21	31	17	69
C	45	34	49	128
D	13	5	20	38
Total	94	96	119	309

TIP MI CONSEJO

Con clasificaciones de dos vías, no probamos hipótesis acerca de probabilidades específicas. Probamos si los dos métodos de clasificación son independientes

Cuando se estudian datos que contienen dos variables, una consideración importante es la *relación entre las dos variables*. ¿La proporción de mediciones en las diversas categorías para el factor 1 depende de cuál categoría del factor 2 se observe? Para el ejemplo de muebles, ¿las proporciones de los diversos defectos varían de turno a turno, o son iguales estas proporciones, independientemente de cuál turno se observe? Recordemos un fenómeno similar llamado *interacción* en el experimento factorial $a \times b$ del capítulo 6. En el análisis de una tabla de contingencia, el objetivo es determinar si un método de clasificación es o no es **contingente** o **dependiente** del otro método de clasificación. Si no lo es, se dice que los dos métodos de clasificación son **independientes**.

PRUEBAS DE CHI CUADRADA PARA LAS DIFERENCIAS ENTRE MÁS DE DOS PROPORCIONES

7.2

Una tabla de contingencia $r \times c$ resulta cuando cada una de las n unidades experimentales se cuenta como si cayera en una de las rc celdas de un experimento multinomial. Cada celda representa un par de niveles de categoría, nivel de fila i y nivel de columna j. A veces, sin embargo, no es aconsejable usar este tipo de diseño experimental, es decir, hacer que n observaciones caigan donde puedan. Por ejemplo, supongamos que se desea estudiar las opiniones de familias estadounidenses acerca de sus niveles de ingreso, es decir, bajos, regulares y altos. Si se seleccionan al azar $n = 1200$ familias para ese estudio, puede que no se encuentre ninguna que se clasifique a sí misma como de bajos ingresos. Podría ser mejor decidir por anticipado hacer un estudio de 400 familias de cada nivel de ingreso. Los datos resultantes aparecerán todavía como clasificación de dos vías, pero los totales de columna se fijan por anticipado.

EJEMPLO 7.2

En otro experimento de prevención de gripe el experimentador decide buscar los registros clínicos de 300 pacientes de cada una de las tres categorías de tratamiento: sin vacuna, una inyección y dos inyecciones. Los $n = 900$ pacientes se encuestarán entonces respecto a su historial de gripe en invierno. El experimento resulta en una tabla de 2×3 con los totales de columna fijos en 300, como se ve en la tabla 7.2. Al fijar los totales de columna, el experimentador ya no tiene un experimento multinomial con $2 \times 3 = 6$ celdas. En cambio, hay tres experimentos binomiales separados, llamémoslos 1, 2 y 3, cada uno con una probabilidad p_j determinada de contraer la gripe y q_j de no contraer la gripe. (Recuerde que para una población binomial, $p_j + q_j = 1$.)

TABLA 7.2 **Casos de gripe para tres tratamientos**

	Sin vacuna	Una inyección	Dos inyecciones	Total
Gripe				r_1
Sin gripe				r_2
Total	300	300	300	n

Supongamos que se utilizó la prueba chi cuadrada para probar la independencia de clasificaciones de fila y columna. Si un tratamiento particular (nivel de columna) no afecta la incidencia de gripe, entonces cada una de las tres poblaciones binomiales debería tener la misma incidencia de gripe para que $p_1 = p_2 = p_3$ y $q_1 = q_2 = q_3$.

La clasificación de 2×3 del ejemplo 7.5 describe una situación en la que la prueba de chi cuadrada de independencia es equivalente a una prueba de la igualdad de $c = 3$ proporciones binomiales. Pruebas de este tipo se llaman **pruebas de homogeneidad** y se usan para comparar diversas poblaciones binomiales. Si hay *más de dos* categorías de renglón con totales de columna fijos, entonces la prueba de independencia es equivalente a una prueba de la igualdad de c conjuntos de proporciones multinomiales.

No es necesario preocuparse de la equivalencia teórica de las pruebas chi cuadrada para estos dos diseños experimentales. Si las columnas (o filas) son fijas o no, el estadístico de prueba se calcula como

$$X^2 = \Sigma \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}} \qquad \text{donde } \hat{E}_{ij} = \frac{r_i c_j}{n}$$

que tiene una distribución chi cuadrada aproximada en muestreo repetido con un $gl = (r - 1)(c - 1)$.

NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

Recuerde el procedimiento general para determinar grados de libertad:

1. Empiece con las rc celdas en la tabla de dos vías.
2. Reste un grado de libertad por cada una de las c poblaciones multinomiales, cuyas probabilidades de columna deben totalizar uno, un total de c *grados de libertad*.
3. Tuvo que estimar $(r - 1)$ probabilidades de fila, pero las probabilidades de columna se fijan por anticipado y no necesitaban ser estimadas. Reste $(r - 1)$.

El total de grados de libertad para la tabla $r \times c$ (columna fija) es

$$rc - c - (r - 1) = rc - c - r + 1 = (r - 1)(c - 1)$$

EJEMPLO 7.3

Una encuesta de opiniones de votantes fue realizada en cuatro distritos políticos del centro de una ciudad, para comparar las fracciones de votantes que están a favor del candidato A. Muestras aleatorias de 200 votantes se encuestaron en cada uno de los cuatro distritos con los resultados que se ven en la tabla 7.3. Los valores entre paréntesis de la tabla son las cantidades de celda esperadas. ¿Los datos presentan suficiente evidencia para indicar que las fracciones de votantes que están a favor del candidato A difieren en los cuatro distritos?

TABLA 7.3 Opiniones de votantes en cuatro distritos

	Distrito				
	1	2	3	4	Total
A favor de A	76 (59)	53 (59)	59 (59)	48 (59)	236
No a favor de A	124 (141)	147 (141)	141 (141)	152 (141)	564
Total	200	200	200	200	800

Solución Como los totales de columna están fijos en 200, el diseño comprende cuatro experimentos binomiales, cada uno de los cuales contiene las respuestas de 200 votantes para cada uno de los cuatro distritos. Para probar la igualdad de las proporciones que están a favor del candidato A en los cuatro distritos, la hipótesis nula

$$H_0: p_1 = p_2 = p_3 = p_4$$

es equivalente a la hipótesis nula

H_0: La proporción a favor del candidato A es independiente del distrito

y será rechazada si el estadístico de prueba X^2 es demasiado grande. El valor observado del estadístico de prueba, $X^2 = 10.722$, y su valor p asociado, .013, se muestran en la figura 7.1. Los resultados son significativos ($P < .025$); esto es, H_0 es rechazada y se puede concluir que hay diferencia en las proporciones de votantes que están a favor del candidato A entre los cuatro distritos.

FIGURA 7.1
Salida *MINITAB* para el ejemplo 7.3

Prueba chi cuadrada: distrito 1, distrito 2, distrito 3, distrito 4

```
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

         Ward 1   Ward 2   Ward 3   Ward 4    Total
1            76       53       59       48      236
          59.00    59.00    59.00    59.00
           4.898    0.610    0.000    2.051

2           124      147      141      152      564
         141.00   141.00   141.00   141.00
           2.050    0.255    0.000    0.858

Total       200      200      200      200      800

Chi-Sq = 10.722 DF = 3, P-Value = 0.013
```

¿Cuál es la naturaleza de las diferencias descubiertas por la prueba chi cuadrada? Para contestar esta pregunta, véase la tabla 7.4, que muestra las proporciones muestrales que están a favor del candidato A en cada uno de los cuatro distritos. Parece que el candidato A está haciéndolo mejor en el primer distrito y peor en el cuarto distrito. ¿Es esto de alguna *significancia práctica* para el candidato? Posiblemente una observación más importante es que el candidato no tiene una pluralidad de votantes en ninguno de los cuatro distritos. Si ésta es una carrera de dos candidatos, el candidato A necesita aumentar su campaña.

TABLA 7.4 **Proporciones a favor del candidato A en cuatro distritos**

Distrito 1	Distrito 2	Distrito 3	Distrito 4
76/200 = .38	53/200 = .27	59/200 = .30	48/200 = .24

7.3 PRUEBA DE INDEPENDENCIA CON CHI CUADRADA

La cuestión de independencia de los dos métodos de clasificación se puede investigar usando una prueba de hipótesis basada en el estadístico chi cuadrada. Éstas son las hipótesis:

H_0: Los dos métodos de clasificación son independientes
H_a: Los dos métodos de clasificación son dependientes

Suponga que denotamos la cantidad de celdas observada en la fila i y la columna j de la tabla de contingencia como O_{ij}. Si se conocieran las cantidades de celda esperadas ($E_{ij} = np_{ij}$) bajo la hipótesis nula de independencia, entonces se podría usar el estadístico chi cuadrada para comparar las cantidades observadas y esperadas. No obstante, los valores esperados no están especificados en H_0, como estaban en los ejemplos previos.

Para explicar cómo estimar estas cantidades de celda esperadas, debemos repasar el concepto de *eventos independientes*. Considere p_{ij}, la probabilidad de que una observación caiga en la fila i y la columna j de la tabla de contingencia. Si las filas y columnas son independientes, entonces

$$\begin{aligned} p_{ij} &= P(\text{observación cae en la fila } i \text{ y columna } j) \\ &= P(\text{observación cae en la fila } i) \times P(\text{observación cae en la columna } j) \\ &= p_i p_j \end{aligned}$$

donde p_i y p_j son las **probabilidades incondicional** o **marginal** de caer en la fila i o columna j, respectivamente. Si se pudiera obtener estimaciones apropiadas de estas probabilidades marginales podrían usarse en lugar de p_{ij} en la fórmula para la cantidad de celda esperada.

Por fortuna, existen estas estimaciones. De hecho, son exactamente lo que en forma intuitiva se elegiría:

- Para estimar la probabilidad de una fila, use

$$\hat{p}_i = \frac{\text{Total de observaciones en fila } i}{\text{número total de observaciones}} = \frac{r_i}{n}$$

- Para estimar la probabilidad de una columna, use

$$\hat{p}_j = \frac{\text{Total de observaciones en la columna } j}{\text{número total de observaciones}} = \frac{c_j}{n}$$

TIP MI CONSEJO

Grados de libertad para una tabla de contingencia $r \times c$: $df = (r-1)(c-1)$

La estimación de la cantidad de celda esperada para la fila i y la columna j se deduce de la suposición de independencia.

CANTIDAD DE CELDA ESPERADA ESTIMADA

$$\hat{E}_{ij} = n\left(\frac{r_i}{n}\right)\left(\frac{c_j}{n}\right) = \frac{r_i c_j}{n}$$

donde r_i es el total para la fila i y c_j es el total para la columna j.

El estadístico de prueba chi cuadrada para una tabla de contingencia con r filas y c columnas se calcula como

$$X^2 = \Sigma \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

y puede mostrarse que tiene una distribución chi cuadrada aproximada con

$$gl = (r-1)(c-1)$$

Si el valor observado de X^2 es demasiado grande, entonces la hipótesis nula de independencia se rechaza.

EJEMPLO 7.4 Consulte el ejemplo 7.1. ¿Los datos presentan suficiente evidencia para indicar que el tipo de defecto de muebles varía con el turno durante el cual la pieza se produjo?

Solución Las cantidades de celda esperadas estimadas se muestran entre paréntesis en la tabla 7.5. Por ejemplo, la cantidad esperada estimada para un defecto tipo C producido durante el segundo turno es

$$\hat{E}_{32} = \frac{r_3 c_2}{n} = \frac{(128)(96)}{309} = 39.77$$

TABLA 7.5 Cantidades de celda esperadas observadas y estimadas

	Turno			
Tipo de defectos	1	2	3	Total
A	15 (22.51)	26 (22.99)	33 (28.50)	74
B	21 (20.99)	31 (21.44)	17 (26.57)	69
C	45 (38.94)	34 (39.77)	49 (49.29)	128
D	13 (11.56)	5 (11.81)	20 (14.63)	38
Total	94	96	119	309

Ahora se pueden usar los valores mostrados en la tabla 7.5 para calcular el estadístico de prueba como

$$X^2 = \Sigma \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

$$= \frac{(15 - 22.51)^2}{22.51} + \frac{(26 - 22.99)^2}{22.99} + \cdots + \frac{(20 - 14.63)^2}{14.63}$$

$$= 19.18$$

Cuando se indiza la distribución chi cuadrada de la tabla 5 del Apéndice con

$$gl = (r - 1)(c - 1) = (4 - 1)(3 - 1) = 6$$

el estadístico de prueba observado es mayor que $\chi^2_{.005} = 18.5476$, lo cual indica que el valor p es menor que .005. Se puede rechazar H_0 y declarar que los resultados son altamente significativos ($P < .005$). Hay suficiente evidencia para indicar que las proporciones de tipos de defecto varían de un turno a otro.

Prueba de independencia chi cuadrada

La siguiente pregunta obvia que se debe formular comprende la naturaleza de la relación entre las dos clasificaciones. ¿Cuál turno produce más de qué tipo de defecto? Al igual que con el experimento factorial del capítulo 6, una vez hallada una dependencia (o interacción) se debe buscar dentro de la tabla en las proporciones relativa o *condicional* para cada nivel de clasificación. Por ejemplo, considere el turno 1, que produjo un total de 94 defectos. Estos defectos se pueden dividir en tipos usando las *proporciones condicionales* para esta muestra que aparecen en la primera columna de la tabla 7.6. Si se sigue el mismo procedimiento para los otros dos turnos, se pueden entonces comparar las distribuciones de tipos de defecto para los tres turnos, como se ve en la tabla 7.6.

Ahora comparemos los tres conjuntos de proporciones (cada uno totaliza 1). Se ve que los turnos 1 y 2 producen defectos en el mismo orden general, tipos C, B, A y D de mayor a menor, aunque en proporciones que difieren. El turno 3 presenta un patrón diferente, casi todos los defectos de tipo C pero seguidos por los tipos A, D y B, en ese orden. Dependiendo de cuál tipo de defecto sea el más importante para el fabricante, cada turno debe ser advertido separadamente acerca de las razones para producir demasiados defectos.

TABLA 7.6 **Probabilidades condicionales para tipos de defecto dentro de tres turnos**

	Turno		
Tipo de defectos	1	2	3
A	$\frac{15}{94} = .16$	$\frac{26}{96} = .27$	$\frac{33}{119} = .28$
B	$\frac{21}{94} = .22$	$\frac{31}{96} = .32$	$\frac{17}{119} = .14$
C	$\frac{45}{94} = .48$	$\frac{34}{96} = .35$	$\frac{49}{119} = .41$
D	$\frac{13}{94} = .14$	$\frac{5}{96} = .05$	$\frac{20}{119} = .17$
Total	1.00	1.00	1.00

NECESITO SABER...

Cómo determinar el número apropiado de grados de libertad

Recuerde el procedimiento general para determinar grados de libertad:

1. Empiece con $k = rc$ categorías o celdas en la tabla de contingencia.
2. Reste un grado de libertad porque todas las probabilidades de celda deben sumar 1.
3. Tuvo que estimar $(r - 1)$ probabilidades de renglón y $(c - 1)$ probabilidades de columna para calcular las cantidades de celda esperadas estimadas. (La última de las probabilidades de fila y de columna está determinada porque las probabilidades *marginales* de fila y columna también totalizan 1.) Reste $(r - 1)$ y $(c - 1)$ grados de libertad (gl).

El total de grados de libertad para la tabla de contingencia $r \times c$ es

$$gl = rc - 1 - (r - 1) - (c - 1) = rc - r - c + 1 = (r - 1)(c - 1)$$

EJEMPLO 7.5

Se realizó un estudio para evaluar la eficacia de una nueva vacuna contra la gripe que había sido administrada en una pequeña comunidad. La vacuna se aplicó sin cargo en una secuencia de dos inyecciones en un periodo de 2 semanas. Algunas personas recibieron la secuencia de dos inyecciones, algunas se presentaron sólo para la primera inyección y otras no recibieron ninguna. Un estudio de mil residentes locales la primavera siguiente dio la información que se ve en la tabla 7.7. ¿Los datos presentan suficiente evidencia para indicar que la vacuna tuvo éxito para reducir el número de casos de gripe en la comunidad?

TABLA 7.7 **Tabla de contingencia de 2 × 3**

	Sin vacuna	Una inyección	Dos inyecciones	Total
Gripe	24	9	13	46
Sin gripe	289	100	565	954
Total	313	109	578	1000

Solución El éxito de la vacuna para reducir el número de casos de gripe se puede evaluar en dos partes:

- Si la vacuna tiene éxito, las proporciones de personas que contraen gripe deberían variar, dependiendo de cuál de los tres tratamientos recibieron.
- Esta dependencia no sólo debe existir, sino que la proporción de personas que contraen gripe debería disminuir a medida que aumenta el tratamiento de prevención de la gripe, de cero a una o dos inyecciones.

La primera parte se puede probar usando la prueba chi cuadrada con estas hipótesis:

H_0: No hay relación entre tratamiento e incidencia de gripe
H_a: La incidencia de gripe depende de la cantidad de tratamiento contra la gripe

TIP MI CONSEJO

Use el valor de X^2 y el valor p de la salida impresa para probar la hipótesis de independencia

Como de costumbre, paquetes de software pueden eliminar todos los tediosos cálculos y, si los datos se introducen correctamente, dan la salida correcta que contiene el valor observado del estadístico de prueba y su valor p. Esa salida impresa, generada por *MINITAB*, se ilustra en la figura 7.2. Se pueden hallar instrucciones para generar esta salida impresa en la sección 7.4. El valor observado del estadístico de prueba, $X^2 = 17.313$, tiene un valor p de .000 y los resultados son declarados altamente significativos. Esto es, la hipótesis nula es rechazada. Hay suficiente evidencia para indicar una relación entre tratamiento e incidencia de la gripe.

FIGURA 7.2
Salida *MINITAB* para el ejemplo 7.5

Prueba chi cuadrada: sin vacuna, una inyección, dos inyecciones

```
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

      No Vaccine One Shot Two Shots   Total
1         24        9         13        46
       14.40     5.01     26.59
       6.404     3.169     6.944

2        289      100       565       954
      298.60   103.99    551.41
       0.309    0.153     0.335

Total    313      109       578      1000

Chi-Sq =  17.313, DF = 2, P-Value = 0.000
```

¿Cuál es la naturaleza de esta relación? Para contestar esta pregunta, véase la tabla 7.8, que da la *incidencia* de gripe en la muestra para cada uno de los tres grupos de tratamiento. La respuesta es obvia. El grupo que recibió dos inyecciones fue menos susceptible a la gripe; una sola inyección no parece reducir la susceptibilidad.

TABLA 7.8 **Incidencia de gripe para tres tratamientos**

Sin vacuna	Una inyección	Dos inyecciones
$\frac{24}{313} = .08$	$\frac{9}{109} = .08$	$\frac{13}{578} = .02$

7.4 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

La prueba chi cuadrada: *Microsoft Excel*

El procedimiento para realizar una prueba de independencia chi cuadrada en *MS Excel* requiere que introduzca tanto los conteos de celda observados como los esperados en una hoja de cálculo de *Excel*. Si se han almacenado en la hoja de cálculo los datos categóricos sin procesar en lugar de los conteos de celda observados, puede ser necesario contar los datos para obtener los conteos de celda antes de continuar.

EJEMPLO 7.6

Suponga que ha registrado el género (M o F) y el nivel universitario (1o. 2o, 3o., 4o., Graduado) para 100 estudiantes de estadística, como se muestra en la siguiente tabla.

	Nivel				
Género	1o.	2o.	3o.	4o.	Graduado
F	16	8	8	8	4
M	9	11	12	12	12

1. Introduzca los valores observados en las primeras cinco columnas de una hoja de cálculo de *Excel*.
2. Calcule (en forma manual) los 10 conteos de celda esperados estimados e introdúzcalos en otro rango en la hoja de cálculo.

3. Coloque el cursor en una celda vacía y use **Formulas ▶ More Functions ▶ Statistical ▶ CHISQ.TEST** (**CHITEST** en versiones anteriores de *Excel*) para generar el cuadro de diálogo que se muestra en la figura 7.3. Resalte o teclee los rangos de celda para los conteos de celda observados y esperados.

FIGURA 7.3

4. Cuando haga clic en **OK**, *MS Excel* calculará el valor *p* asociado con la prueba de independencia chi cuadrada. Para estos datos, el valor *p* grande (.153) indica un resultado no significativo. Hay evidencia insuficiente para indicar que el género de un estudiante depende de su nivel en clase.

NOTA: *MS Excel* no proporciona un comando **único que permita** realizar la prueba de bondad del ajuste chi cuadrada; sin embargo, podría crear en forma manual fórmulas en *MS Excel* para ejecutar esta prueba y obtener el valor *p* apropiado.

La prueba chi cuadrada: *MINITAB*

Existen varios procedimientos en el paquete *MINITAB* para analizar datos categóricos. El procedimiento apropiado depende de si los datos representan una clasificación de una vía (un solo experimento multinomial) o una clasificación de dos vías o tabla de contingencia. Si los *datos categóricos sin elaborar* se han guardado en la hoja de cálculo *MINITAB* más que las *cantidades de celda observadas*, puede ser necesario totalizar o clasificar en forma cruzada los datos para obtener las cantidades de celda antes de continuar.

EJEMPLO 7.7

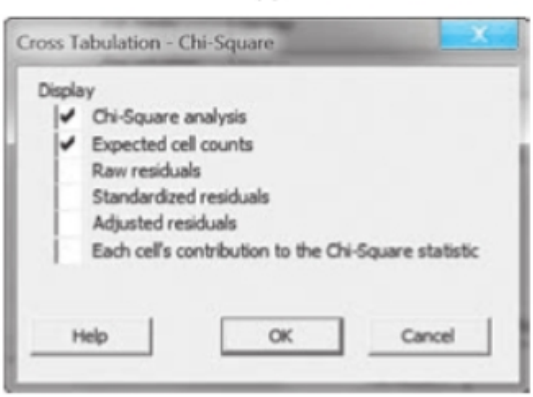
Suponga que ha registrado el género (M o F) y el nivel universitario (1o., 2o., 3o., 4o., Graduado) para 100 estudiantes de estadística. La hoja de cálculo de *MINITAB* contendría dos columnas de 100 observaciones cada una. Cada fila contendría un género del individuo en la columna 1 y el nivel universitario en la columna 2.

1. Para obtener los conteos de celda observados (O_{ij}) para la tabla de contingencia de 2 × 5, use **Stat ▶ Tables ▶ Cross Tabulation and Chi-Square** para generar el cuadro de diálogo que se muestra en la figura 7.4a.

FIGURA 7.4

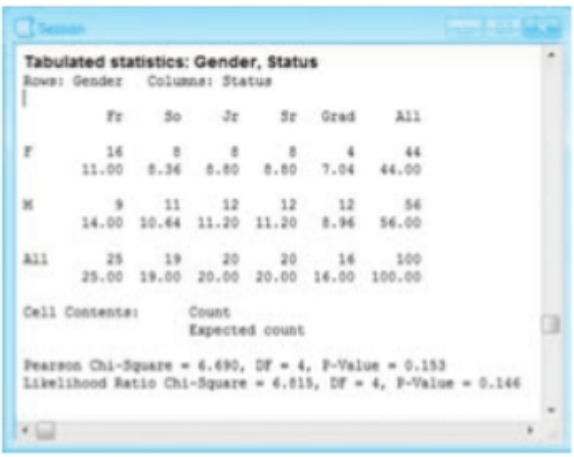
(a)

(b)



2. Bajo "Categorical Variables", seleccione "Género" para la variable de fila y "Nivel" para la variable de columna. Deje en blanco los cuadros marcados "For Layers" y "Frequencies are in:". Asegúrese que está seleccionado el cuadro etiquetado "Display Counts".
3. Haga clic en el botón **Chi-Square…** para mostrar el cuadro de diálogo que aparece en la figura 7.4b. Seleccione los cuadros para "Chi-Square Analysis" y "Expected Cell Counts". Haga clic en **OK** dos veces. Esta secuencia de comandos no sólo tabula la tabla de contingencia sino también realiza la prueba de independencia chi cuadrada y muestra los resultados en la ventana Session mostrada en la figura 7.5. Para los datos de género/nivel universitario, el valor p grande ($P = .153$) indica un resultado no significativo. Hay evidencia insuficiente para indicar que el género de un estudiante depende de su nivel en clase.

FIGURA 7.5



```
Tabulated statistics: Gender, Status
Rows: Gender   Columns: Status

          Fr     So     Jr     Sr   Grad    All

F         16      8      8      8      4     44
       11.00   8.36   8.80   8.80   7.04  44.00

M          9     11     12     12     12     56
       14.00  10.64  11.20  11.20   8.96  56.00

All       25     19     20     20     16    100
       25.00  19.00  20.00  20.00  16.00 100.00

Cell Contents:      Count
                    Expected count

Pearson Chi-Square = 6.690, DF = 4, P-Value = 0.153
Likelihood Ratio Chi-Square = 6.815, DF = 4, P-Value = 0.146
```

4. Si los conteos de celda observados en la tabla de contingencia ya han sido tabulados, simplemente introduzca los conteos en c columnas de la hoja de cálculo de *MINITAB*, use **Stat ▶ Tables ▶ Chi-Square Test (Two-Way Table in Worksheet)** y seleccione las columnas apropiadas antes de hacer clic en **OK**. Para los datos de género/nivel universitario, puede introducir los conteos en las columnas C3–C7 como se muestra en la figura 7.6. La salida resultante será etiquetada de manera diferente pero se verá exactamente como la salida en la figura 7.5.

FIGURA 7.6

EJEMPLO 7.8

Se puede establecer una prueba simple de un experimento multinomial único al considerar si las proporciones de hombres y mujeres estudiantes de estadística son iguales; es decir, $p_1 = .5$ y $p_2 = .5$.

1. En *MINITAB 15* o *16*, use **Stat ▶ Tables ▶ Chi-Square Goodness-of-Fit Test (One Variable)** para desplegar el cuadro de diálogo que se muestra en la figura 7.7. Si tiene datos categóricos sin procesar en una columna, haga clic en el botón "Categorical data:" e introduzca la columna "Género" en la celda. Si tiene valores de resumen de conteos observados para cada categoría, elija "Observed counts". Luego introduzca la columna que contiene los conteos observados o teclee los conteos observados para cada categoría.

FIGURA 7.7

2. Para esta prueba, se puede elegir "Equal proportions" para probar H_0: $p_1 = p_2 = .5$. Cuando tiene diferentes proporciones para cada categoría, use "Specific proportions". Puede almacenar las proporciones para cada categoría en una columna, elija "Input column" e introduzca la columna. Si desea teclear la proporción para cada categoría, seleccione "Input constants" y teclee las proporciones para las categorías correspondientes. Haga clic en **OK**.
3. La salida resultante incluirá varias gráficas junto con los valores para O_i y E_i para cada categoría, el valor observado del estadístico de prueba, $X^2 = 1.44$ y su valor $p = 0.230$, el cual no es significativo. Hay evidencia insuficiente para indicar una diferencia en la proporción de hombres y mujeres estudiantes de estadística.

NOTA: Si usa una versión previa de *MINITAB* tendrá que determinar los conteos de celda observados y esperados, e introducirlos en columnas separadas en la hoja de cálculo. Luego use **Calc ▶ Calculator** y la expresión **SUM(('O' - 'E')**2/'E')** para calcular el valor observado del estadístico de prueba.

EJERCICIOS

TÉCNICAS BÁSICAS

7.1 Muestras aleatorias de 200 observaciones se seleccionaron de cada una de tres poblaciones y cada observación se clasificó de acuerdo con si cayó en una de tres categorías mutuamente exclusivas:

	Categoría			
Población	1	2	3	Total
1	108	52	40	200
2	87	51	62	200
3	112	39	49	200

Se desea saber si los datos dan suficiente evidencia para indicar que las proporciones de observaciones en las tres categorías dependen de la población de la cual se sacaron.

a. Dé el valor de X^2 para la prueba.

b. Dé la región de rechazo para la prueba para $\alpha = .01$.

c. Exprese sus conclusiones.

d. Encuentre el valor p aproximado para la prueba e interprete su valor.

APLICACIONES

7.2 Pollos enfermos Se piensa que una enfermedad particular en pollos no es comunicable. Para probar esta teoría, 30000 pollos se dividieron al azar en tres grupos de 10000. Un grupo no tenía contacto con pollos enfermos, otro tenía contacto moderado y el tercero tenía contacto frecuente. Después de seis meses se recabaron datos sobre el número de pollos enfermos en cada grupo de 10000. ¿Los datos dan suficiente evidencia para indicar una dependencia entre la cantidad de contacto entre aves enfermas y no enfermas y la incidencia de la enfermedad? Use $\alpha = .05$.

	Sin contacto	Contacto moderado	Contacto frecuente
Enfermedad	87	89	124
No enfermedad	9913	9911	9876
Total	10000	10000	10000

7.3 Evolución: ¿A favor o en contra? Según una encuesta del Pew Research Center, 55% de los adultos jóvenes (con edades de 18–29 años) creen que la evolución es la mejor explicación para el desarrollo de la vida humana.[1] Cuando los datos se clasifican además de acuerdo con si los que respondieron tenían una afiliación religiosa o no, esta proporción cambió para aquellos que no tenían una afiliación religiosa. Los datos que siguen reflejan los resultados de esta encuesta.

	Afiliación religiosa	No afiliado	Total
Sí	47	152	199
No	53	98	151
Total	100	250	350

a. ¿Los datos indican que la proporción de adultos jóvenes que creen que la evolución proporciona la mejor respuesta para el desarrollo de la vida humana difiere para aquellos con una afiliación religiosa contra aquellos sin una? Use $\alpha = .05$.

b. Si existen diferencias significativas, explique cuáles cambios parecen haber tenido lugar cuando se incluye la afiliación religiosa en la clasificación.

7.4 Suponga que un estudio de consumidores compendia las respuestas de $n = 307$ personas en una tabla de contingencia de tres filas y cinco columnas. ¿Cuántos grados de libertad están asociados con el estadístico de prueba chi cuadrada?

DATOS EX1423

7.5 Color de cabello El color del cabello y de los ojos que sigue se obtuvo por autoinforme de una muestra de estadounidenses caucásicos nacidos entre 1957 y 1965 (en la actualidad de 45–53 años de edad).[2] Los siguientes datos se adaptaron de ese estudio.

	Rubio claro	Rubio	Castaño claro	Castaño	Negro	Rojo	Total
Hombres	4	46	45	176	23	10	304
Mujeres	5	58	69	164	12	12	320

a. ¿Hay suficiente evidencia para concluir que la proporción de individuos con estos colores de cabello difieren para hombres y mujeres? Use $\alpha = .05$.

b. ¿Hay alguna celda con un número esperado menor a 5? De ser así, combine esas celdas con aquellas que están junto a ellas y vuelva a analizar los datos. ¿Difieren los resultados finales?

DATOS EX1425

7.6 Trabajo a distancia Como alternativa al horario flexible, numerosas compañías permiten que sus empleados hagan parte de su trabajo en casa. Las personas de una muestra aleatoria de 300 trabajadores se clasificaron de acuerdo con el salario y número de días de trabajo por semana que pasan en casa.

	Días de trabajo en casa por semana		
Salario	Menos de uno	Al menos uno, pero no todos	Todos en casa
Menos de $25 000	38	16	14
$25 000 a $49 999	54	26	12
$50 000 a $74 999	35	22	9
Más de $75 000	33	29	12

a. ¿Los datos presentan suficiente evidencia para indicar que el salario depende del número de días de trabajo que pasan en casa? Pruebe usando $\alpha = .05$.

b. Use la tabla 5 del Apéndice para aproximar el valor p para esta prueba de hipótesis. ¿El valor p confirma sus conclusiones de la parte **a**?

7.7 Accidentes fatales Se analizaron datos de accidentes para determinar los números de accidentes fatales para automóviles de tres tamaños. Los datos de 346 accidentes son como sigue:

	Pequeño	Mediano	Grande
Fatal	67	26	16
No fatal	128	63	46

¿Los datos indican que la frecuencia de accidentes fatales depende del tamaño de los automóviles? Escriba un breve párrafo que describa sus resultados estadísticos y las implicaciones prácticas de éstos.

7.8 Color y forma de una flor Un botánico realiza un cruce secundario de petunias con factores independientes que controlan la forma de la hoja y el color de la flor, donde el factor A representa el color rojo, a representa color blanco, B representa hojas redondas y b representa hojas largas. De acuerdo con el modelo de Mendel, las plantas deben exhibir las características AB, Ab, aB

y *ab* en la proporción 9:3:3:1. De 160 plantas experimentales, se observaron los números siguientes:

AB	*Ab*	*aB*	*ab*
95	30	28	7

¿Hay suficiente evidencia para refutar el modelo de Mendel al nivel $\alpha = .01$?

7.9 Ajuste de una distribución binomial Se puede usar una prueba de bondad de ajuste para determinar si en realidad todos los criterios para un experimento binomial se han satisfecho en una aplicación determinada. Suponga que un experimento consistente en cuatro intentos se repitió cien veces. El número de repeticiones en las que se obtuvo un número dado de éxitos se registró en la tabla:

Resultados posibles (número de éxitos)	Número de veces obtenido
0	11
1	17
2	42
3	21
4	9

Estime p (suponiendo que el experimento fue binomial), obtenga estimaciones de las frecuencias de celda esperadas y pruebe la bondad del ajuste. Para determinar el número apropiado de grados de libertad para X^2, observe que p fue estimado por una combinación lineal de las frecuencias observadas.

7.10 Antibióticos e infección A veces ocurren infecciones cuando se aplican transfusiones de sangre durante operaciones quirúrgicas. Se realizó un experimento para determinar si la inyección de anticuerpos redujo la probabilidad de infección. Un examen de los registros de 138 pacientes produjo los datos que se ven en la tabla siguiente. ¿Los datos dan suficiente evidencia para indicar que inyecciones de anticuerpos afectan la probabilidad de infecciones en transfusiones? Pruebe usando $\alpha = .05$.

	Infección	No infección
Anticuerpos	4	78
Sin anticuerpos	11	45

DATOS EX1453

7.11 ¿Sus alimentos son seguros? ¿Cuánta confianza tiene de que es seguro comer los alimentos que compra? Esta pregunta se hizo en una encuesta de *CBS News*.[3] Los datos que siguen reflejan los resultados de las respuestas a esta encuesta.

	Mucha confianza	Alguna confianza	No demasiada confianza	Ninguna confianza en absoluto	Total
Hombres	210	241	68	5	524
Mujeres	129	306	73	16	524
Total	329	547	141	21	1048

a. ¿Hay suficiente evidencia para concluir que hay diferencias significativas en las respuestas entre hombres y mujeres en el nivel de significancia de $\alpha = .05$?

b. Encuentre el valor p aproximado para la prueba.

DATOS EX1456

7.12 Medicina que sabe bien Pfizer Canada Inc. es una empresa farmacéutica que hace acitromicina, antibiótico en suspensión con sabor a cereza, que se usa para tratar infecciones bacterianas en niños. Para comparar el sabor de su producto con tres medicamentos de la competencia, Pfizer probó 50 niños sanos y 20 adultos sanos. Entre otras medidas para probar el sabor, registraron el número de probadores que calificaron cada una de las cuatro suspensiones de antibiótico como el de mejor sabor.[4] Los resultados se muestran en la tabla siguiente. ¿Hay diferencia en la percepción del mejor sabor entre adultos y niños? Si es así, ¿cuál es la naturaleza de la diferencia y por qué es de importancia práctica para la compañía farmacéutica?

	Sabor de antibiótico			
	Plátano	Cereza*	Fruta silvestre	Fresa-Plátano
Niños	14	20	7	9
Adultos	4	14	0	2

* Acitromicina producida por Pfizer Canada Inc.

7.13 Pescar un resfriado ¿La probabilidad de contraer un resfriado está influida por el número de contactos sociales que usted tenga? Un estudio de Sheldon Cohen, profesor de psicología de la Universidad Carnegie Mellon, parece mostrar que cuantas más relaciones sociales tenga alguien *menos susceptible* es a los resfriados.[5] Un grupo de 276 hombres y mujeres sanos se agruparon de acuerdo con su número de relaciones (padres, amigos, miembros de la iglesia, vecinos). Se les expone entonces a un virus que causa los resfriados. Una adaptación de los resultados se muestra en la tabla.

	Número de relaciones		
	Tres o menos	Cuatro o cinco	Seis o más
Resfriado	49	57	34
No resfriado	31	43	62
Total	80	100	96

a. ¿Los datos dan suficiente evidencia para indicar que la susceptibilidad a los resfriados es afectada por el número de relaciones que se tenga? Pruebe al nivel de significancia de 5%.

b. Con base en los resultados de la parte **a**, describa la naturaleza de la relación entre las dos variables categóricas: incidencia de resfriado y número de relaciones sociales. ¿Sus observaciones concuerdan con las conclusiones del autor?

Fuentes

1. "Religion Among the Millennials: Less Religiously Active Than Older Americans, But Fairly Traditional In Other Ways," http://pewforum.org/Age/Religion-Among-the-Millennials.aspx, 17 de febrero de 2010.
2. Jason Malloy, "Gene Expression: NLSY Blogging: Eye and Hair Color of Americans," http://www.gnxp.com/blog/2008/12/nlsy-blogging-eye-and-hair-color-of.php, 31 de diciembre de 2008.
3. "Food Safety," *CBS News Poll*, www.pollingreport.com, 17–22, diciembre de 2009.
4. Doreen Matsui, R. Lim, T. Tschen y M. J. Rieder, "Assessment of the Palatability of b-Lactamase-Resistant Antibiotics in Children," *Archives of Pediatric Adolescent Medicine* 151 (Junio de 1997):599.
5. Adaptado de David L. Wheeler, "More Social Roles Means Fewer Colds," *Chronicle of Higher Education* XLIII, no. 44 (Julio 11 de 1997):A13.

8

Regresión lineal y correlación

AP Photo/David Zalubowski

OBJETIVOS GENERALES

En este capítulo consideramos la situación en la que el valor medio de una variable aleatoria y está relacionado con otra variable x. Al medir tanto y como x para cada unidad experimental, con lo cual se generan datos bivariados, se puede usar la información dada por x para estimar el valor promedio de y y para predecir valores de y para valores de x asignados previamente.

ÍNDICE DEL CAPÍTULO

- Análisis de correlación
- Análisis de varianza para regresión lineal
- El método de mínimos cuadrados
- Estimación y predicción con uso de la recta ajustada
- Prueba de la utilidad del modelo de regresión lineal: inferencias acerca de β, la prueba F de ANOVA, y r^2
- Un modelo probabilístico lineal sencillo

NECESITO SABER...

Cómo estar seguro de que mis cálculos son correctos

8.1 DIAGRAMA DE DISPERSIÓN, EL MODELO LINEAL, REGRESIÓN SIMPLE

Los estudiantes de último año de preparatoria, los de primer año de universidad, sus padres, así como la administración de una universidad, están preocupados por el rendimiento académico de un estudiante después de haberse inscrito en la universidad. ¿Se puede estimar o predecir el promedio de calificaciones (GPA; siglas de *grade point average*) de un estudiante al terminar su último año de preparatoria, antes de inscribirse en la universidad? A primera vista éste podría ser un problema difícil aunque se espera que los estudiantes altamente motivados, que se hayan graduado con calificaciones altas de una preparatoria, alcancen un GPA alto cuando terminen el primer año. Por otra parte, no es probable que los que carezcan de motivación o que hayan obtenido sólo un éxito moderado en preparatoria salgan bien. Se esperaría que el rendimiento académico de un estudiante sea una función de diversas variables:

- Rango en su grupo de preparatoria
- Nivel general de preparatoria
- Alto promedio GPA
- Calificaciones del SAT

Este problema es de naturaleza más bien general. Usted se interesará por una variable aleatoria y (promedio GPA) relacionada con diversas variables independientes. El objetivo es crear una *ecuación de predicción* que exprese y como función de estas variables independientes. A continuación, si se pueden medir las variables independientes, es posible sustituir estos valores en la ecuación de predicción y obtener la predicción para y, es decir, el promedio GPA del estudiante en nuestro ejemplo. Pero, ¿cuáles variables deben usarse para hacer la predicción? ¿Qué tan fuerte es su relación con y? ¿Cómo se construye una buena ecuación de predicción para y como función de las variables seleccionadas para la predicción? Estas interrogantes se analizan en los capítulos 7 y 9.

En este capítulo restringimos nuestra atención al sencillo problema de predecir y como función lineal de una sola variable x de pronóstico. Recuerde que usamos la ecuación de una recta para describir la relación entre x y y, y describimos la fuerza de la relación usando el coeficiente de correlación r. Nos apoyaremos en algunos de estos resultados cuando repasemos el tema de regresión lineal y correlación.

Considere el problema de tratar de predecir el valor de una respuesta y basada en el valor de una variable independiente x.

$$y = a + bx$$

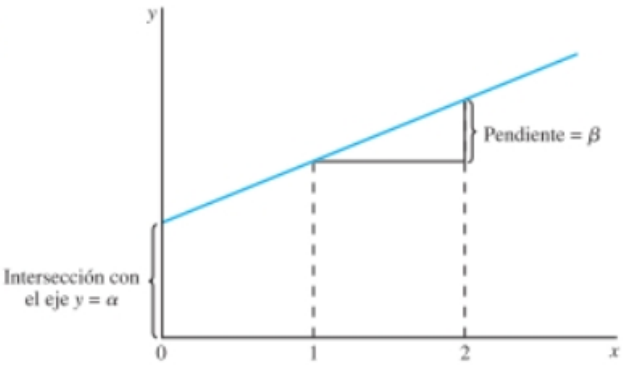
estuvo basada en una *muestra* de n observaciones bivariadas tomadas de una *población* más grande de medidas. La recta que describe la relación entre y y x en la *población* es semejante a la recta de mejor ajuste de la *muestra*, pero no es igual. ¿Cómo se puede construir un **modelo de población** para describir la relación entre una variable aleatoria y y una variable independiente x relacionada?

Se empieza por suponer que la variable de interés, y, está *linealmente* relacionada a una variable independiente x. Para describir la relación lineal, se puede usar el **modelo determinista**

$$y = \alpha + \beta x$$

donde α es la intersección y, es decir, el valor de y cuando $x = 0$, y β es la pendiente de la recta, definida como el cambio en y para un cambio unitario en x, como se muestra en la figura 8.1. Este modelo describe una relación determinista entre la variable de interés y, a veces llamada **variable de respuesta**, y la variable independiente x, denominada **variable de pronóstico**. Esto es, la ecuación lineal determina un valor exacto de y cuando se da el valor de x. ¿Este modelo es realista para una situación experimental? Considere el siguiente ejemplo.

FIGURA 8.1
Intersección con el eje y y pendiente de una recta



TIP MI CONSEJO

Pendiente = cambio en y para un cambio unitario en x

intersección con eje y = valor de y cuando $x = 0$

MI APPLET EN LÍNEA

Construir una gráfica de dispersión

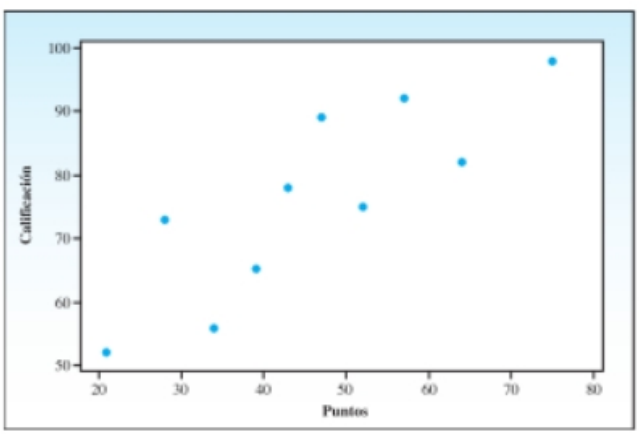
La tabla 8.1 muestra las calificaciones del examen de aprovechamiento en matemáticas de una muestra aleatoria de $n = 10$ estudiantes de primer año de universidad, junto con sus calificaciones finales en cálculo. Una gráfica bivariada de estos puntos y calificaciones se da en la figura 8.2. Observe que los puntos *no están exactamente sobre una recta* sino que más bien parecen ser desviaciones alrededor de una recta subyacente. Una forma sencilla de modificar el modelo determinista es agregar un **componente aleatorio de error** para explicar las desviaciones de los puntos alrededor de la recta. Una respuesta particular y se describe usando el **modelo probabilístico**

$$y = \alpha + \beta x + \epsilon$$

TABLA 8.1
Calificaciones de examen de aprovechamiento en matemáticas y puntuaciones finales en cálculo para estudiantes de primer año de universidad

Estudiante	Puntuación en el examen de aprovechamiento en matemáticas	Calificación final en cálculo
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

FIGURA 8.2
Gráfica de dispersión de datos de la tabla 8.1



La primera parte de la ecuación, $\alpha + \beta x$, llamada **recta de medias**, describe el valor promedio de y para un valor determinado de x. El componente de error ϵ permite que cada respuesta individual y se desvíe de la recta de medias en una pequeña cantidad.

Para usar este *modelo probabilístico* para hacer inferencias es necesario ser más específico acerca de esta "pequeña cantidad", ϵ.

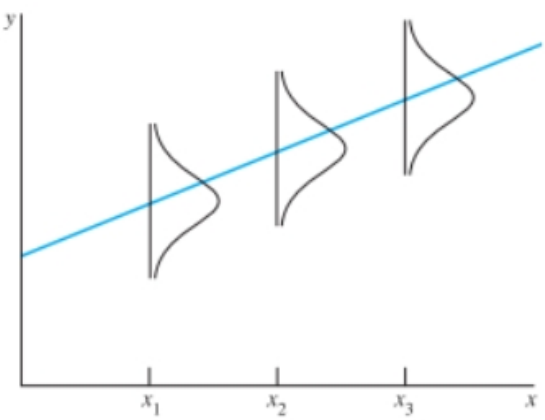
SUPOSICIONES ACERCA DEL ERROR ALEATORIO ϵ

Suponga que los valores de ϵ satisfacen estas condiciones:

- Son independientes en el sentido probabilístico
- Tienen una media de 0 y una varianza común igual a σ^2
- Tienen una distribución normal de probabilidad

Estas suposiciones acerca del error aleatorio ϵ se muestran en la figura 8.3 para tres valores fijos de x, por ejemplo x_1, x_2 y x_3. Observe la similitud entre estas suposiciones y las suposiciones necesarias para las pruebas en los capítulos 4 y 6. Repasaremos estas suposiciones más adelante en este capítulo y daremos algunas herramientas de diagnóstico para que usted las use en la verificación de su validez.

FIGURA 8.3
Modelo probabilístico lineal



Recuerde que este modelo se ha creado para una población de mediciones que por lo general es desconocida, pero puede usar información muestral para estimar los valores de α y β, que son los coeficientes de la recta de medias, $E(y) = \alpha + \beta x$. Estas estimaciones se usan para formar la recta de mejor ajuste para un conjunto de datos determinado, llamado **recta de mínimos cuadrados** o **recta de regresión**. En la siguiente sección repasamos la forma de calcular el punto de intersección y la pendiente de esta recta.

8.2 MÉTODOS DE AJUSTE DE DATOS, MÉTODOS DE MÍNIMOS CUADRADOS PARA AJUSTAR UNA RECTA DE REGRESIÓN, PROPIEDADES DE LOS ESTIMADORES

El procedimiento estadístico para hallar la recta de mejor ajuste para un conjunto de datos bivariados hace, matemáticamente, lo que en forma visual se realiza cuando se mueve una regla hasta que se hayan reducido al mínimo las distancias verticales o desviaciones, de la regla a un conjunto de puntos. La fórmula de la recta de mejor ajuste es

$$\hat{y} = a + bx$$

donde a y b son las estimaciones de los parámetros α y β de intersección y pendiente, respectivamente. La recta ajustada para los datos de la tabla 8.1 se muestra en la figura 8.4. Las rectas

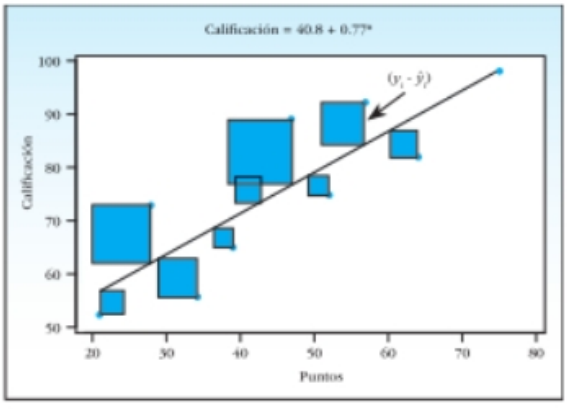
TIP MI CONSEJO

Pendiente = coeficiente de x

Intersección con eje y = término constante

verticales trazadas desde la recta de predicción $(x_i, \hat{y}_i)$ hasta cada punto (x_i, y_i) representan las desviaciones de los puntos desde la recta; es decir, $(y_i - \hat{y}_i)$.

FIGURA 8.4
Método de mínimos cuadrados



APPLET EN LÍNEA
Método de mínimos cuadrados

Observe que algunos puntos están debajo de la recta de predicción, y por tanto $(y_i - \hat{y}_i)$ será negativo. Para evitar que las distancias positivas y negativas se "cancelen entre sí", elegimos minimizar las distancias desde los puntos hasta la recta ajustada, usando el **principio de mínimos cuadrados**.

PRINCIPIO DE MÍNIMOS CUADRADOS

La recta que reduce al mínimo la suma de cuadrados de las desviaciones de los valores observados de y desde los pronosticados es la **recta de mejor ajuste.** La suma del cuadrado de las desviaciones por lo general se denomina **suma de cuadrados de error** (SSE) y se define como

$$\text{SSE} = \Sigma(y_i - \hat{y}_i)^2 = \Sigma(y_i - a - bx_i)^2$$

Observe la recta de regresión y los puntos de datos de la figura 8.4. SSE es la suma del cuadrado de las distancias representada por el área de los cuadros azules.

Hallar los valores de a y b, las estimaciones de α y β, usa cálculo diferencial, que está fuera del propósito de este libro. En lugar de derivar sus valores, simplemente presentaremos fórmulas para calcular los valores de a y b, llamados **estimadores de mínimos cuadrados** de α y β. Usaremos una notación que está basada en **las sumas de cuadrados** para las variables del problema de regresión, que es semejante en forma a las sumas de cuadrados empleadas en el capítulo 6.

Usted debe usar el método de entrada de datos para su calculadora científica para introducir los datos muestrales.

- Si su calculadora tiene sólo una función estadística de una variable, todavía puede ahorrar tiempo al hallar las sumas necesarias y sumas de cuadrados.
- Si su calculadora tiene una función estadística de dos variables o si tiene una calculadora graficadora, la calculadora en forma automática guarda todas las sumas y sumas de cuadrados, así como los valores de a, b y el coeficiente de correlación r.
- Asegúrese de consultar el manual de su calculadora para hallar la forma más fácil de obtener los estimadores de mínimos cuadrados.

ESTIMADORES DE MÍNIMOS CUADRADOS DE α y β

$$b = \frac{S_{xy}}{S_{xx}} \quad \text{y} \quad a = \bar{y} - b\bar{x}$$

donde las cantidades S_{xy} y S_{xx} están definidas como

$$S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n}$$

y

$$S_{xx} = \Sigma(x_i - \bar{x})^2 = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

EJEMPLO 8.1 Encuentre la recta de predicción de mínimos cuadrados para los datos de la calificación en cálculo de la tabla 8.1.

Solución Use los datos de la tabla 8.2 y el método de introducción de datos en su calculadora científica para hallar las siguientes sumas de cuadrados:

$$S_{xx} = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n} = 23\,634 - \frac{(460)^2}{10} = 2474$$

$$S_{xy} = \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n} = 36\,854 - \frac{(460)(760)}{10} = 1894$$

$$\bar{y} = \frac{\Sigma y_i}{n} = \frac{760}{10} = 76 \qquad \bar{x} = \frac{\Sigma x_i}{n} = \frac{460}{10} = 46$$

TABLA 8.2 Cálculos para los datos de la tabla 8.1

	y_i	x_i	x_i^2	$x_i y_i$	y_i^2
	65	39	1521	2535	4225
	78	43	1849	3354	6084
	52	21	441	1092	2704
	82	64	4096	5248	6724
	92	57	3249	5244	8464
	89	47	2209	4183	7921
	73	28	784	2044	5329
	98	75	5625	7350	9604
	56	34	1156	1904	3136
	75	52	2704	3900	5625
Suma	760	460	23 634	36 854	59 816

Entonces

$$b = \frac{S_{xy}}{S_{xx}} = \frac{1894}{2474} = .76556 \quad \text{y} \quad a = \bar{y} - b\bar{x} = 76 - (.76556)(46) = 40.78424$$

TIP MI CONSEJO

Se puede predecir y para un valor determinado de x al sustituir x en la ecuación para hallar $\hat{y}$

La recta de regresión de mínimos cuadrados es entonces

$$\hat{y} = a + bx = 40.78424 + .76556x$$

La gráfica de esta recta se ve en la figura 8.4. Ahora se puede usar para predecir y para un valor determinado de x, ya sea consultando la figura 8.4 o sustituyendo el valor apropiado de x en la ecuación. Por ejemplo, si un alumno de primer año obtuvo $x = 50$ en el examen, la calificación pronosticada de cálculo del estudiante es (usando precisión completa de decimales)

$$\hat{y} = a + b(50) = 40.78424 + (.76556)(50) = 79.06$$

NECESITO SABER...

Cómo estar seguro de que mis cálculos son correctos

- Tenga cuidado con los errores de redondeo. Lleve al menos seis cifras significativas y haga redondeo sólo al informar el resultado final.
- Use una calculadora científica o graficadora para hacer todo el trabajo. Casi todas las calculadoras estimarán los valores de a y b si se le introducen correctamente los datos.
- Use un programa de computadora si tiene acceso a ella.
- Siempre grafique los datos y la recta. Si la recta no se ajusta a los puntos, ¡es probable que el usuario tenga un error!

Análisis de varianza para regresión lineal

En el capítulo 6 utilizamos los procedimientos del análisis de varianza para dividir la variación total del experimento en partes atribuidas a diversos factores de interés para el experimentador. En un análisis de regresión, la respuesta y está relacionada con la variable independiente x. En consecuencia, la variación total de la variable de respuesta y, dada por

$$\text{TSS} = S_{yy} = \Sigma(y_i - \bar{y})^2 = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}$$

está dividida en dos partes:

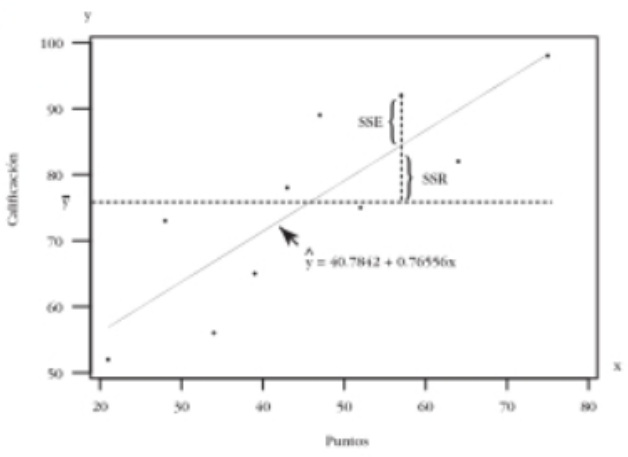
- La SSR (suma de cuadrados para regresión) mide la cantidad de variación explicada al usar la recta de regresión con una variable independiente x
- La SSE (suma de cuadrados de error) mide la variación "residual" en los datos que no está explicada por la variable independiente x

de modo que

$$\text{TSS} = \text{SSR} + \text{SSE}$$

Para un valor particular de la respuesta y_i, se puede visualizar este desglose en la variación usando las distancias verticales ilustradas en la figura 8.5. Se puede ver que la SSR es la suma del cuadrado de desviaciones de las diferencias entre la respuesta estimada sin usar $x(\bar{y})$ y la respuesta estimada usando x (la recta de regresión, $\hat{y}$); la SSE es la suma del cuadrado de diferencias entre la recta de regresión ($\hat{y}$) y el punto y.

FIGURA 8.5
Desviaciones desde la recta ajustada



No es demasiado difícil demostrar algebraicamente que

$$\text{SSR} = \Sigma(\hat{y}_i - \bar{y})^2 = \Sigma(a + bx_i - \bar{y})^2 = \Sigma(\bar{y} - b\bar{x} + bx_i - \bar{y})^2 = b^2\Sigma(x_i - \bar{x})^2$$
$$= \left(\frac{S_{xy}}{S_{xx}}\right)^2 S_{xx} = \frac{(S_{xy})^2}{S_{xx}}$$

Como TSS = SSR + SSE, se puede completar la partición al calcular

$$\text{SSE} = \text{TSS} - \text{SSR} = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

Recuerde del capítulo 6 que cada una de las diversas fuentes de variación, cuando se dividen entre los **grados de libertad** apropiados, da una estimación de la variación en el experimento. Estas estimaciones se denominan **medias cuadráticas**, MS = SS/*gl*, y se ven en una tabla ANOVA.

Al examinar los grados de libertad asociados con cada una de estas sumas de cuadrados, observe que el total de grados de libertad para n mediciones es $(n - 1)$. Como la estimación de la recta de regresión, $\hat{y} = a + bx_i = \bar{y} - b\bar{x} + bx_i$, abarca la estimación de *un* parámetro *adicional* β, hay *un* grado de libertad asociado con la SSR, dejando $(n - 2)$ grados de libertad con la SSE.

Al igual que con todas las tablas ANOVA que hemos estudiado, el error medio cuadrático

$$\text{MSE} = s^2 = \frac{\text{SSE}}{n - 2}$$

es un estimador insesgado de la varianza subyacente σ^2. La tabla de análisis de varianza se muestra en la tabla 8.3.

TABLA 8.3 **Análisis de varianza para regresión lineal**

Fuente	*gl*	SS	MS
Regresión	1	$\frac{(S_{xy})^2}{S_{xx}}$	MSR
Error	$n - 2$	$S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$	MSE
Total	$n - 1$	S_{yy}	

Para los datos de la tabla 8.1, se puede calcular

$$\text{TSS} = S_{yy} = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n} = 59\,816 - \frac{(760)^2}{10} = 2\,056$$

$$\text{SSR} = \frac{(S_{xy})^2}{S_{xx}} = \frac{(1\,894)^2}{2\,474} = 1\,449.9741$$

de modo que

$$\text{SSE} = \text{TSS} - \text{SSR} = 2\,056 - 1\,449.9741 = 606.0259$$

y

$$\text{MSE} = \frac{\text{SSE}}{n - 2} = \frac{606.0259}{8} = 75.7532$$

La tabla de análisis de varianza, parte de la *salida de regresión lineal* generada por *MINITAB*, es la sección inferior sombreada de la salida de computadora de la figura 8.6a. Las primeras dos rectas dan la ecuación de la recta de mínimos cuadrados, $\hat{y} = 40.8 + .766x$. Las estimaciones de mínimos cuadrados a y b están dadas con mayor precisión en la columna marcada "Coef". La salida de *MS Excel* para los mismos datos se muestra en la figura 8.6b. La tabla

FIGURA 8.6a
Salida *MINITAB* para los datos de la tabla 8.1

Análisis de regresión: *y* contra *x*

```
The regression equation is
y = 40.8 + 0.766 x

Predictor        Coef      SE Coef          T        P
Constant       40.784        8.507       4.79    0.001
x              0.7656       0.1750       4.38    0.002

S = 8.70363        R-Sq = 70.5%     R-Sq(adj) = 66.8%

Analysis of Variance
Source           DF          SS          MS          F        P
Regression        1      1450.0      1450.0      19.14    0.002
Residual Error    8       606.0        75.8
Total             9      2056.0
```

FIGURA 8.6b
Salida de *MS Excel* para los datos de la tabla 8.1

SUMMARY OUTPUT							
Regression Statistics							
Multiple R	0.8398						
R Square	0.7052						
Adjusted R Square	0.6684						
Standard Error	8.7036						
Observations	10						
ANOVA							
	df	*SS*	*MS*	*F*	*Significance F*		
Regression	1	1449.974	1449.974	19.141	0.002		
Residual	8	606.026	75.753				
Total	9	2056					
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*
Intercept	40.784	8.507	4.794	0.001	21.167	60.401	21.167
Score	0.766	0.175	4.375	0.002	0.362	1.169	0.362

de ANOVA está en medio de la salida sombreada; las estimaciones de mínimos cuadrados se encuentran en la parte inferior de la salida sombreada en la columna etiquetada "Coefficients". Se pueden hallar instrucciones para generar esta salida impresa en la sección 8.7.

Las salidas de computadora también dan alguna información acerca de la variación en el experimento. Cada una de las estimaciones de mínimos cuadrados, a y b, tiene un error estándar asociado, marcado "SE Coef" en la figura 8.6a y "Standard Error" en la figura 8.6b. Hacia la mitad de la salida *MINITAB* se encuentra la mejor estimación insesgada de σ —$s = \sqrt{\text{MSE}} = \sqrt{75.7532} = 8.70363$— que mide el **error residual**, la variación no explicada o "sobrante" del experimento. Esta misma medida se encuentra en la porción superior de la salida de *MS Excel* etiquetada "Standard Error". No sorprenderá saber que los estadísticos t y F y sus valores p hallados en la salida impresa se usan para probar hipótesis estadísticas.

TIP MI CONSEJO

Busque *a* y *b* en la columna llamada "Coef" o "Coefficients"

8.3 INTERVALOS DE CONFIANZA

Otra forma de hacer inferencias acerca del valor de β es construir un intervalo de confianza para β y examinar el rango de posibles valores para β.

UN INTERVALO DE CONFIANZA $(1 - \alpha)100\%$ PARA β

$$b \pm t_{\alpha/2}(\text{SE})$$

donde $t_{\alpha/2}$ está basada en $(n - 2)$ grados de libertad, $s^2 = \text{MSE}$, y

$$\text{SE} = \sqrt{\frac{s^2}{S_{xx}}} = \sqrt{\frac{\text{MSE}}{S_{xx}}}$$

EJEMPLO 8.2

Encuentre una estimación de intervalo de confianza de 95% de la pendiente β para los datos de las calificaciones en cálculo de la tabla 8.1.

Solución Sustituyendo valores previamente calculados en

$$b \pm t_{.025}\sqrt{\frac{\text{MSE}}{S_{xx}}}$$

tendremos

$$.766 \pm 2.306\sqrt{\frac{75.7532}{2474}}$$

$$.766 \pm .404$$

El intervalo de confianza de 95% resultante es .362 a 1.170. Como el intervalo no contiene 0, se puede concluir que el verdadero valor de β no es 0 y se puede rechazar la hipótesis nula H_0: $\beta = 0$ a favor de H_a: $\beta \neq 0$, conclusión que está de acuerdo con los hallazgos del ejemplo 8.2. Además, la estimación del intervalo de confianza indica que hay un aumento desde sólo .4 hasta 1.2 puntos en una puntuación de examen de cálculo por cada aumento de 1 punto en la puntuación del examen de aprovechamiento.

Si usted utiliza un programa de cómputo para hacer un análisis de regresión, encontrará la estadística t y su valor p en la salida impresa. En la segunda sección de la salida de *MINITAB* en la figura 8.7a, encontrará la estimación b de mínimos cuadrados de la recta de la pendiente marcada "x", junto con su error estándar "SE Coef", el valor calculado del estadístico de prueba "T" usado para probar la hipótesis H_0: $\beta = 0$ y su valor p, "P". Encontrará la misma información en la última línea de la salida de *MS Excel* en la figura 8.7b, junto con los límites de confianza superior e inferior de un intervalo de confianza de 95% para β. La prueba t para regresión significativa, H_0: $\beta = 0$, tiene un valor p de $P = .002$, y la hipótesis nula se rechaza, como en el ejemplo 8.3. Hay una relación lineal significativa entre x y y.

FIGURA 8.7a
Salida *MINITAB* para los datos de calificaciones en cálculo

Análisis de regresión: *y* contra *x*

```
The regression equation is
y = 40.8 + 0.766 x

Predictor      Coef     SE Coef          T          P
Constant     40.784       8.507       4.79      0.001
x            0.7656      0.1750       4.38      0.002

S = 8.70363    R-Sq = 70.5%    R-Sq(adj) = 66.8%

Analysis of Variance
Source            DF        SS        MS        F        P
Regression         1    1450.0    1450.0    19.14    0.002
Residual Error     8     606.0      75.8
Total              9    2056.0
```

TIP MI CONSEJO

Busque el error estándar de *b* en la columna marcada "SE Coef" en la salida de *MINITAB* y "Standard Error" en la salida de *MS Excel*

FIGURA 8.7b
Salida de *MS Excel* para los datos de las calificaciones de cálculo

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.8398
R Square	0.7052
Adjusted R Square	0.6684
Standard Error	8.7036
Observations	10

ANOVA

	df	*SS*	*MS*	*F*	*F de significancia*
Regression	1	1449.974	1449.974	19.141	0.002
Residual	8	606.026	75.753		
Total	9	2056			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	40.784	8.507	4.794	0.001	21.167	60.401
x	0.766	0.175	4.375	0.002	0.362	1.169

8.4 PRUEBAS DE HIPÓTESIS

¿La recta de regresión de mínimos cuadrados es útil? Es decir, ¿la ecuación de regresión que utiliza información dada por x es considerablemente mejor que la pronosticadora simple $\bar{y}$ que no se apoya en x? Si la variable independiente x *no es útil* en el modelo de población $y = \alpha + \beta x + \epsilon$, entonces el valor de y no cambia para valores diferentes de x. La única forma en que esto ocurre para todos los valores de x es cuando la pendiente β de la recta de medias es igual a 0. Esto indicaría que la relación entre y y x no es lineal, de modo que la pregunta inicial acerca de la utilidad de la variable independiente x se puede expresar también como: ¿Hay una relación lineal entre x y y?

Se puede contestar esta pregunta usando ya sea una prueba de hipótesis o un intervalo de confianza para β. Estos procedimientos están basados en la distribución muestral de b, el estimador muestral de la pendiente β. Se puede demostrar que si las suposiciones acerca del error aleatorio ϵ son válidas, entonces el estimador b tiene una distribución normal en muestreo repetido con media

$$E(b) = \beta$$

y error estándar dado por

$$\text{SE} = \sqrt{\frac{\sigma^2}{S_{xx}}}$$

donde σ^2 es la varianza del error aleatorio ϵ. Como el valor de σ^2 se estima con $s^2 = \text{MSE}$, se pueden basar inferencias en la estadística dada por

$$t = \frac{b - \beta}{\sqrt{\text{MSE}/S_{xx}}}$$

que tiene una distribución t con $gl = (n - 2)$, los grados de libertad asociados con MSE.

PRUEBA DE HIPÓTESIS RESPECTO A LA PENDIENTE DE UNA RECTA

1. Hipótesis nula: H_0: $\beta = \beta_0$
2. Hipótesis alternativa

Prueba de una cola	**Prueba de dos colas**
H_a: $\beta > \beta_0$ (o $\beta < \beta_0$)	H_a: $\beta \neq \beta_0$

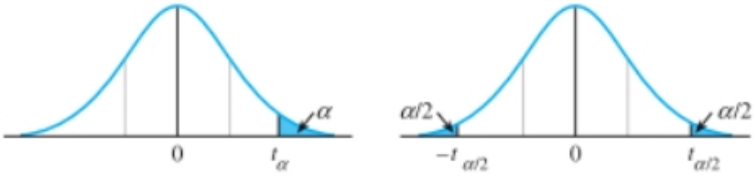
3. Estadístico de prueba $t = \dfrac{b - \beta_0}{\sqrt{\text{MSE}/S_{xx}}}$

   Cuando se satisfacen las suposiciones dadas en la sección 8.1, el estadístico de prueba tendrá una distribución t de Student con $(n - 2)$ grados de libertad.

4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	**Prueba de dos colas**
$t > t_\alpha$ (o $t < t_\alpha$ cuando la hipótesis alternativa sea H_a: $\beta > \beta_0$)	$t > t_{\alpha/2}$ o $t < -t_{\alpha/2}$

o cuando valor $p < \alpha$



Los valores de t_α y $t_{\alpha/2}$ correspondientes a $(n - 2)$ grados de libertad pueden encontrarse usando la tabla 4 del Apéndice.

EJEMPLO 8.3

Determine si hay una relación lineal significativa entre las calificaciones en cálculo y las puntuaciones de examen de la tabla 8.1. Pruebe al nivel de significancia de 5%.

Solución Las hipótesis a probar son

$$H_0: \beta = 0 \quad \text{contra} \quad H_a: \beta \neq 0$$

APPLET EN LÍNEA

Prueba *t* para la pendiente y probabilidades de Student

y el valor observado del estadístico de prueba se calcula como

$$t = \frac{b - 0}{\sqrt{\text{MSE}/S_{xx}}} = \frac{.7656 - 0}{\sqrt{75.7532/2474}} = 4.38$$

con $(n - 2) = 8$ grados de libertad. Con $\alpha = .05$, se puede rechazar H_0 cuando $t > 2.306$ o $t < -2.306$. Como el valor observado del estadístico de prueba cae en la región de rechazo, H_0 es rechazada y se puede concluir que hay una relación lineal significativa entre las calificaciones en cálculo y la puntuación de examen para la población de estudiantes de primer año de universidad.

El análisis de varianza de la prueba *F*

TIP MI CONSEJO

Las pruebas *F* de ANOVA siempre son de una cola (cola superior)

La parte del análisis de varianza de la salida impresa de las figuras 8.7a y 8.7b muestra un estadístico F dado por

$$F = \frac{\text{MSR}}{\text{MSE}} = 19.14$$

con grado de libertad 1 en el numerador y $(n - 2) = 8$ grados de libertad en el denominador. Esto es una *estadística equivalente de prueba* que también se puede usar para probar la hipótesis H_0: $\beta = 0$. Observe que, dentro del error de redondeo, el valor de F es igual a t^2 con valor p idéntico. En este caso, si se usa una precisión de cinco lugares decimales antes de redondeo, se encuentra que $t^2 = (.76556/.17498)^2 = (4.37513)^2 = 19.14175 \approx 19.14 = F$ como se da en la salida impresa. Esto no es por casualidad y resulta del hecho de que el cuadrado de un estadístico t con gl grados de libertad tiene la misma distribución que un estadístico F con grados de libertad 1 en el numerador y gl en el denominador. La prueba F es una prueba más general de la utilidad del modelo y se puede usar cuando el modelo tenga más de una variable independiente.

Fuerza de la relación: el coeficiente de determinación

¿Qué tan bien se ajusta el modelo de regresión? Para contestar esta pregunta, se puede usar una medida relacionada con el *coeficiente de correlación r*. Recuerde que

$$r = \frac{s_{xy}}{s_x s_y} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} \qquad \text{para} -1 \leq r \leq 1$$

donde s_{xy}, s_x y s_y se definieron en el capítulo 3 y las diversas sumas de cuadrados se definieron en la sección 8.2.

La suma de cuadrados para regresión, SSR, en el análisis de varianza, mide la parte de la variación total TSS $= S_{yy}$, que puede ser explicada por la regresión de y en x. La parte restante, SSE, es la variación "no explicada" atribuida al error aleatorio. Una forma de medir la fuerza de la relación entre la variable de respuesta y y la variable de predicción x es calcular el **coeficiente de determinación**, la proporción de la variación total que es explicada por la regresión lineal de y en x. Para los datos de calificaciones en cálculo, esta proporción es igual a

TIP MI CONSEJO

En las salidas impresas de computadora, r^2 a menudo se da como un **porcentaje** más que como una proporción

$$\frac{\text{SSR}}{\text{TSS}} = \frac{1450}{2056} = .705 \quad \text{o} \quad 70.5\%$$

Puesto que TSS $= S_{yy}$ y SSR $= \dfrac{(S_{xy})^2}{S_{xx}}$, se puede escribir

$$\frac{\text{SSR}}{\text{TSS}} = \frac{(S_{xy})^2}{S_{xx}S_{yy}} = \left(\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}\right)^2 = r^2$$

Por tanto, el coeficiente de determinación, que se calculó como SSR/TSS, es simplemente el cuadrado del coeficiente de correlación r. Es la entrada marcada "R-Sq" en la figura 8.7a y "R cuadrada" en la figura 8.7b.

Recuerde que la tabla del análisis de varianza aísla la variación debida a regresión (SSR) de la variación total del experimento. Al hacer esto se reduce la cantidad de *variación aleatoria* del experimento, ahora medida por SSE en lugar de TSS. En este contexto, el **coeficiente de determinación**, r^2, se puede definir como sigue:

Definición El **coeficiente de determinación** r^2 se puede interpretar como el porcentaje de reducción en la variación total en el experimento obtenido al usar la recta de regresión $\hat{y} = a + bx$ en lugar de ignorar x y usar la media muestral $\bar{y}$ para predecir la variable de respuesta y.

TIP MI CONSEJO

r^2 se denomina "R-Sq" en la salida impresa de *MINITAB* y "R Square" en la salida impresa de *Excel*

Para los datos de calificaciones en cálculo, una reducción de $r^2 = .705$ o sea 70.5% es considerable. El modelo de regresión funciona muy bien.

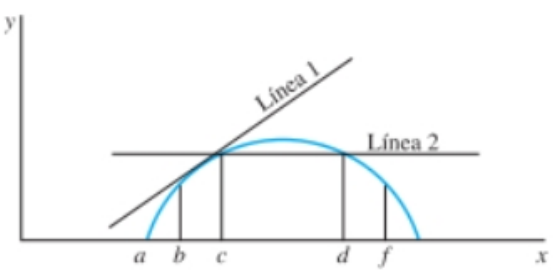
Interpretación de los resultados de una regresión significativa

Una vez que usted haya efectuado la prueba t o la prueba F para determinar la significancia de la regresión lineal, con todo cuidado debe interpretar sus resultados. La pendiente β de la recta de medias se estima con base en datos de sólo una región de observación en particular. Incluso si no rechaza la hipótesis nula de que la pendiente de la recta es igual a 0, no necesariamente significa que y y x no estén relacionadas. Puede ser que haya cometido un error tipo II, declarando falsamente que la pendiente es 0 y que x y y no están relacionadas.

Ajuste del modelo erróneo

Puede ocurrir que y y x estén perfectamente relacionadas en una forma no lineal, como se ve en la figura 8.8. A continuación veamos tres posibilidades:

FIGURA 8.8
Relación curvilínea



- Si se tomaron observaciones sólo dentro del intervalo $b < x < c$, la relación aparecería lineal con pendiente positiva.
- Si se tomaron observaciones sólo dentro del intervalo $d < x < f$, la relación aparecería lineal con pendiente negativa.
- Si se tomaron observaciones sobre el intervalo $c < x < d$, la recta estaría ajustada con una pendiente cercana a 0, lo cual indica que no hay relación lineal entre y y x.

Para el ejemplo que se ilustra en la figura 8.8 ninguna recta describe con precisión la verdadera relación entre x y y, que es en realidad una *relación curvilínea.* En este caso, hemos escogido un *modelo erróneo* para describir la relación. A veces este tipo de error se puede detectar usando gráficas residuales, que es el tema de la sección 8.6.

Extrapolación

TIP MI CONSEJO

Es peligroso tratar de predecir valores de y fuera del rango de los datos ajustados

Un problema serio es aplicar los resultados de un análisis de regresión lineal a valores de x que *no estén incluidos* dentro del rango de los datos ajustados. Esto se llama **extrapolación** y puede conducir a errores graves en la predicción, como se ve para la línea 1 de la figura 8.8. Los resultados de una predicción serían buenos en el intervalo $b < x < c$ pero sobreestimarían gravemente los valores de y para $x > c$.

Causalidad

Cuando haya una regresión significativa de y y x, es tentador concluir que x *causa a* y. No obstante, es posible que una o más variables desconocidas que ni siquiera se hayan medido y que no estén incluidas en el análisis puedan estar causando la relación observada. En general, el estadístico informa los resultados de un análisis pero deja las conclusiones respecto a la causalidad a científicos e investigadores que son expertos en estos campos de actividad. Estos expertos están mejor preparados para tomar esas decisiones.

8.5 CORRELACIÓN

El coeficiente de correlación, r, formalmente denominado **coeficiente de correlación muestral producto-momento de Pearson,** se define a continuación.

COEFICIENTE DE CORRELACIÓN PRODUCTO-MOMENTO DE PEARSON

$$r = \frac{s_{xy}}{s_x s_y} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} \quad \text{para } -1 \le r \le 1$$

TIP MI CONSEJO

r está siempre entre -1 y +1

Las varianzas y covarianza se pueden hallar por cálculo directo, con el uso de una calculadora con capacidad de estadística de dos variables o usando un paquete de estadística como el *MINITAB* o *MS Excel*. Las varianzas y covarianza se calculan como

$$s_{xy} = \frac{S_{xy}}{n-1} \qquad s_x^2 = \frac{S_{xx}}{n-1} \qquad s_y^2 = \frac{S_{yy}}{n-1}$$

y use S_{xy}, S_{xx} y S_{yy}, las mismas cantidades empleadas antes en el análisis de regresión en este capítulo. En general, cuando una muestra de n individuos o unidades experimentales se selecciona y dos variables se miden en cada individuo o unidad, de modo que *ambas variables sean aleatorias,* el coeficiente de correlación r es la medida apropiada de linealidad para usar en esta situación.

EJEMPLO 8.4

Las estaturas y los pesos de $n = 10$ jugadores atacantes de futbol se seleccionan al azar de un equipo de estrellas de un condado. Calcule el coeficiente de correlación para las estaturas (en pulgadas) y pesos (en libras) dado en la tabla 8.4.

TABLA 8.4 **Estaturas y pesos de $n = 10$ estrellas profundos**

Jugador	Estatura, x	Peso, y
1	73	185
2	71	175
3	75	200
4	72	210
5	72	190
6	75	195
7	67	150
8	69	170
9	71	180
10	69	175

Solución Debe usarse el método apropiado de entrada de datos de su calculadora científica para verificar los cálculos para las sumas de cuadrados y productos cruz,

$$S_{xy} = 328 \qquad S_{xx} = 60.4 \qquad S_{yy} = 2610$$

usando las fórmulas de cálculo dadas antes en este capítulo. Entonces

$$r = \frac{328}{\sqrt{(60.4)(2610)}} = .8261$$

o sea $r = .83$. Este valor de r es bastante cercano a 1, el máximo valor posible de r, que indica una relación lineal positiva bastante fuerte entre estatura y peso.

TIP MI CONSEJO

El signo de r es siempre **igual** al signo de la pendiente b

Hay una relación directa entre las fórmulas de cálculo para el coeficiente de correlación r y la pendiente de la recta de regresión b. Como el numerador de ambas cantidades es S_{xy}, tanto r como b tienen el mismo signo. Por tanto, el coeficiente de correlación tiene estas propiedades generales:

- Cuando $r = 0$, la pendiente es $b = 0$ y no hay relación lineal entre x y y.
- Cuando r es positiva, b también es positiva y hay una relación lineal positiva entre x y y.
- Cuando r es negativa, b también es negativa y hay una relación lineal negativa entre x y y.

En la sección 8.4 demostramos que

$$r^2 = \frac{\text{SSR}}{\text{TSS}} = \frac{\text{TSS} - \text{SSE}}{\text{TSS}}$$

En esta forma se puede ver que r^2 nunca puede ser mayor que 1, de modo que $-1 \le r \le 1$. Además, se observa la relación entre la variación aleatoria (medida por SSE) y r^2.

MI APPLET EN LÍNEA

Explorar la correlación

- Si no hay variación aleatoria y todos los puntos caen en la recta de regresión, entonces $\text{SSE} = 0$ y $r^2 = 1$.
- Si los puntos están dispersos en forma aleatoria y no hay variación explicada por regresión, entonces $\text{SSR} = 0$ y $r^2 = 0$.

La figura 8.9 muestra cuatro gráficas de dispersión típicas y sus coeficientes de correlación asociados. Observe que en la gráfica de dispersión (d) parece haber una relación curvilínea entre x y y, pero r es aproximadamente 0, lo cual refuerza el hecho de que r es una medida de una relación *lineal* (no *curvilínea*) entre dos variables.

FIGURA 8.9
Algunas gráficas de dispersión típicas con valores aproximados de r

Considere una población generada al medir dos variables aleatorias en cada unidad experimental. En esta población *bivariada*, el **coeficiente de correlación poblacional** ρ se calcula e interpreta como está en la muestra. En esta situación, el experimentador puede probar la hipótesis de que no hay correlación entre las variables x y y usando una estadística de prueba que sea *exactamente equivalente* a la prueba de la pendiente β de la sección 8.4. El procedimiento de prueba se muestra a continuación.

PRUEBA DE HIPÓTESIS RESPECTO AL COEFICIENTE DE CORRELACIÓN ρ

1. Hipótesis nula: H_0: $\rho = 0$
2. Hipótesis alternativa:

Prueba de una cola	**Prueba de dos colas**
H_a: $\rho > 0$ (o $\rho < 0$)	H_a: $\rho \neq 0$

3. Estadístico de prueba: $t = r\sqrt{\dfrac{n-2}{1-r^2}}$

   Cuando se satisfacen las suposiciones dadas en la sección 8.1, el estadístico de prueba tendrá una distribución t de Student con $(n-2)$ grados de libertad.
4. Región de rechazo: rechazar H_0 cuando

Prueba de una cola	**Prueba de dos colas**
$t > t_\alpha$ (o $t < -t_\alpha$ cuando la hipótesis alternativa sea H_a: $\rho < 0$)	$t > t_{\alpha/2}$ o $t < -t_{\alpha/2}$

   o cuando valor $p < \alpha$

Los valores de t_α y $t_{\alpha/2}$ correspondientes a $(n-2)$ grados de libertad pueden encontrarse usando la tabla 4 en el Apéndice.

TIP MI CONSEJO

Se puede demostrar que

$$t = r\sqrt{\frac{n-2}{1-r^2}} = \frac{b-0}{\sqrt{\text{MSE}/S_{xx}}}$$

EJEMPLO 8.5 Consulte los datos de estatura y peso del ejemplo 8.4. La correlación entre estatura y peso se calculó como $r = .8261$. ¿Esta correlación es significativamente diferente de 0?

Solución Para probar las hipótesis

$$H_0: \rho = 0 \quad \text{contra} \quad H_a: \rho \neq 0$$

el valor del estadístico de prueba es

$$t = r\sqrt{\frac{n-2}{1-r^2}} = .8261\sqrt{\frac{10-2}{1-(.8261)^2}} = 4.15$$

que para $n = 10$ tiene una distribución t con 8 grados de libertad. Como este valor es mayor que $t_{.005} = 3.355$, el valor p de dos colas es menor que $2(.005) = .01$ y la correlación se declara significativa al nivel de 1% ($P < .01$). El valor $r^2 = .8261^2 = .6824$ significa que alrededor de 68% de la variación en una de las variables es explicada por la otra. La salida impresa *MINITAB* de la figura 8.10 muestra la correlación r y el valor p exacto para probar su significancia.

TIP MI CONSEJO

El valor t y valor p para probar H_0: $\rho = 0$ será idéntico al valor t y valor p para probar H_0: $\beta = 0$

FIGURA 8.10
Salida *MINITAB* para el ejemplo 8.5

Correlaciones: x, y

```
Pearson correlation of x and y = 0.826
P-Value = 0.003
```

Si los coeficientes lineales de correlación entre y y cada una de las dos variables x_1 y x_2 se calculan como .4 y .5, respectivamente, no se concluye que un medio de predicción que use ambas variables sea $[(.4)^2 + (.5)^2] = .41$, o sea una reducción de 41% en la suma de cuadrados de desviaciones. En realidad, x_1 y x_2 podrían estar bastante correlacionados y por tanto contribuir prácticamente con la misma información para la predicción de y.

Por último, recuerde que r es una medida de **correlación lineal** y que x y y podrían estar perfectamente relacionadas por alguna función **no lineal** cuando el valor observado de r sea igual a 0. El problema de estimar o predecir y usando información dada por varias variables independientes, $x_1, x_2, \ldots, x_k$, es el tema del capítulo 9.

8.6 PREDICCIÓN

Ahora que usted ha

- probado y ajustado la recta de regresión, $\hat{y} = a + bx$, para asegurarse que es útil para predicción y
- empleado las herramientas de diagnóstico para asegurarse que ninguna de las suposiciones de regresión han sido violadas

está listo para usar la recta para uno de sus dos propósitos:

- Estimar el valor promedio de y para un valor determinado de x
- Predecir un valor particular de y para un valor determinado de x

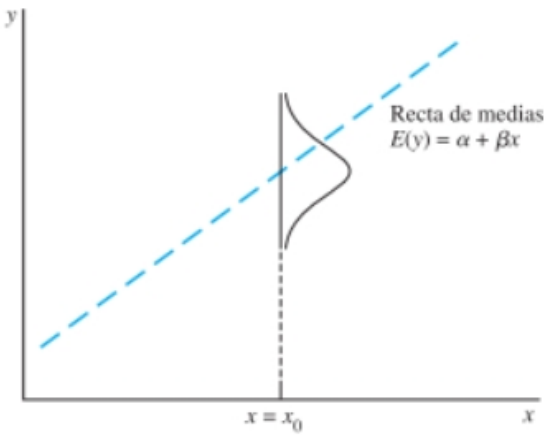
La muestra de n pares de observaciones ha sido seleccionada de una población en la que el valor *promedio* de y está relacionado con el valor de la variable de pronóstico x por la **recta de medias**,

$$E(y) = \alpha + \beta x$$

una recta desconocida, que se muestra como recta interrumpida en la figura 8.11. Recuerde que para un valor fijo de x, por ejemplo x_0, los valores *particulares* de y se desvían desde la

recta de medias. Estos valores de y se supone que tienen una distribución normal con media igual a $\alpha + \beta x_0$ y varianza σ^2, como se ve en la figura 8.11.

FIGURA 8.11
Distribución de y para $x = x_0$



Como los valores calculados de a y b varían de una muestra a otra, cada nueva muestra produce una diferente recta de regresión $\hat{y} = a + bx$, que se puede usar ya sea para estimar la recta de medias o para predecir un valor particular de y. La figura 8.12 muestra una de las posibles configuraciones de la recta ajustada (azul), la recta de medias desconocida (gris), y un valor particular de y (el punto azul).

FIGURA 8.12
Error al estimar $E(y)$ y predecir y

¿A qué distancia estará nuestro estimador $\hat{y} = a + bx_0$ desde la cantidad a estimar o predecir? Esto depende, como siempre, de la variabilidad de nuestro estimador, medido por su **error estándar**. Se puede demostrar que

$$\hat{y} = a + bx_0$$

el valor estimado de y cuando $x = x_0$, es un estimador insesgado de la recta de medias, $\alpha + \beta x_0$ y que $\hat{y}$ está normalmente distribuida con el error estándar de $\hat{y}$ estimado por

$$\text{SE}(\hat{y}) = \sqrt{\text{MSE}\left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right)}$$

La estimación y prueba están basadas en la estadística

$$t = \frac{\hat{y} - E(y)}{\text{SE}(\hat{y})}$$

que tiene una distribución t con $(n - 2)$ grados de libertad.

Para formar un intervalo de confianza $(1 - \alpha)100\%$ para el valor promedio de y cuando $x = x_0$, medido por la recta de medias, $\alpha + \beta x_0$, se puede usar la forma usual para un intervalo de confianza basado en la distribución t:

$$\hat{y} \pm t_{\alpha/2}\text{SE}(\hat{y})$$

TIP MI CONSEJO

Para un valor determinado de x, el intervalo de predicción es siempre más ancho que el intervalo de confianza

No obstante, si se escoge predecir un valor *particular* de y cuando $x = x_0$, hay algún error adicional en la predicción debido a la desviación de y desde la recta de medias. Si examinamos la figura 8.12, se puede ver que el error en predicción tiene dos componentes:

- El error al usar la recta ajustada para estimar la recta de medias
- El error causado por la desviación de y desde la recta de medias, medida por σ^2

La varianza de la diferencia entre y y $\hat{y}$ es la suma de estas dos varianzas y forma la base para el error estándar de $(y - \hat{y})$ empleado para predicción:

$$\text{SE}(y - \hat{y}) = \sqrt{\text{MSE}\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}$$

y el intervalo de predicción $(1 - \alpha)100\%$ se forma como

$$\hat{y} \pm t_{\alpha/2}\text{SE}(y - \hat{y})$$

INTERVALOS DE CONFIANZA Y PREDICCIÓN $(1 - \alpha)100\%$

- Para estimar el valor promedio de y cuando $x = x_0$:

$$\hat{y} \pm t_{\alpha/2}\sqrt{\text{MSE}\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}$$

- Para predecir un valor particular de y cuando $x = x_0$:

$$\hat{y} \pm t_{\alpha/2}\sqrt{\text{MSE}\left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}$$

donde $t_{\alpha/2}$ es el valor de t con $(n - 2)$ grados de libertad y área $\alpha/2$ a su derecha.

EJEMPLO 8.6

Use la información del ejemplo 8.1 para estimar el promedio de calificaciones en cálculo para estudiantes cuya puntuación de aprovechamiento es 50, con un intervalo de confianza de 95%.

Solución La estimación puntual de $E(y|x_0 = 50)$, el promedio de calificación en cálculo para estudiantes cuya puntuación de aprovechamiento es 50, es

$$\hat{y} = 40.78424 + .76556(50) = 79.06$$

El error estándar de $\hat{y}$ es

$$\sqrt{\text{MSE}\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]} = \sqrt{75.7532\left[\frac{1}{10} + \frac{(50 - 46)^2}{2474}\right]} = 2.840$$

y el intervalo de confianza de 95% es

$$79.06 \pm 2.306(2.840)$$
$$79.06 \pm 6.55$$

Nuestros resultados indican que el promedio de calificación en cálculo para estudiantes cuya puntuación es de 50 en el examen de aprovechamiento estará entre 72.51 y 85.61.

EJEMPLO 8.7

Un estudiante tomó el examen de aprovechamiento y obtuvo 50 pero todavía no ha tomado el examen de cálculo. Usando la información del ejemplo 8.1, prediga la calificación en cálculo para este estudiante, con un intervalo de predicción de 95%.

Solución El valor predicho de y es $\hat{y} = 79.06$, como en el ejemplo 8.6. No obstante, el error en predicción se mide con $SE(y - \hat{y})$, y el intervalo de predicción de 95% es

$$79.06 \pm 2.306\sqrt{75.7532\left[1 + \frac{1}{10} + \frac{(50 - 46)^2}{2474}\right]}$$

$$79.06 \pm 2.306(9.155)$$

$$79.06 \pm 21.11$$

o sea de 57.95 a 100.17. El intervalo de predicción es *más ancho* que el intervalo de confianza del ejemplo 8.6 por la variabilidad extra al predecir el valor real de la respuesta y.

Es frecuente que un punto particular sobre la recta de medias sea de interés para experimentadores, α, la intersección con el eje y, es decir, el valor promedio de y cuando $x_0 = 0$.

EJEMPLO 8.8

Antes de ajustar una recta a los datos de la calificación en cálculo y las puntuaciones de aprovechamiento, se puede pensar que una puntuación de 0 en el examen de aprovechamiento puede predecir una calificación de 0 en el examen de cálculo. Esto implica que debemos ajustar un modelo con α igual a 0. ¿Los datos apoyan la hipótesis de un punto de intersección en 0?

Solución Se puede contestar esta pregunta al construir un intervalo de confianza de 95% para el punto α de intersección con el eje y, que es el valor promedio de y cuando $x = 0$. La estimación de α es

$$\hat{y} = 40.784 + .76556(0) = 40.784 = a$$

y el intervalo de confianza de 95% es

$$\hat{y} \pm t_{\alpha/2}\sqrt{\text{MSE}\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}\right]}$$

$$40.784 \pm 2.306\sqrt{75.7532\left[\frac{1}{10} + \frac{(0 - 46)^2}{2474}\right]}$$

$$40.784 \pm 19.617$$

o de 21.167 a 60.401, un intervalo que no contiene el valor $\alpha = 0$. En consecuencia, es improbable que la intersección con el eje y sea 0. Se debe incluir un punto de intersección diferente de cero en el modelo $y = \alpha + \beta x + \epsilon$.

Para esta situación especial en la que estamos interesados en probar o estimar el punto de intersección α con el eje y para la recta de medias, las inferencias comprenden la estimación muestral a. La prueba para un punto de intersección en 0 se da en la figura 8.13 en la recta sombreada marcada como "Constante". El coeficiente dado como 40.784 es a, con error estándar dado en la columna marcada "SE Coef " como 8.507, que concuerda con el valor calculado en el ejemplo 8.8. El valor de $t = 4.79$ se encuentra al dividir a entre su error estándar con valor $p = .001$.

FIGURA 8.13
Parte de la salida *MINITAB* para el ejemplo 8.8

Predictor	Coef	SE Coef	T	P
Constant	40.784	8.507	4.79	0.001
x	0.7656	0.1750	4.38	0.002

Se puede ver que es bastante lento calcular manualmente estos intervalos de predicción y estimación, además de que es difícil mantener precisión en los cálculos, pero por fortuna hay

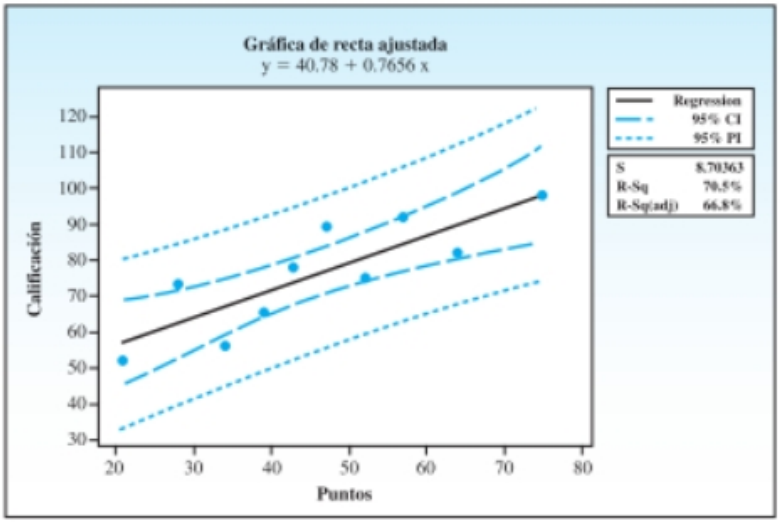
programas de cómputo que pueden hacer los cálculos. El comando de regresión *MINITAB* tiene una opción para estimación o predicción cuando se especifica el valor (o valores) necesario(s) de x. La salida impresa de la figura 8.14 da los valores de $\hat{y} = 79.06$ marcados "Fit", el error estándar de **[símbolo]**, SE($\hat{y}$), marcado "SE Fit", el *intervalo de confianza* para el valor promedio de y cuando $x = 50$, marcado "95.0% CI", y el intervalo de predicción para y cuando $x = 50$, marcado "95.0% PI".

FIGURA 8.14
Opción *MINITAB* para estimación y predicción

```
Predicted Values for New Observations
New Obs     Fit     SE Fit        95% CI             95% PI
1         79.06       2.84    (72.51,   85.61)   (57.95,  100.17)

Values of Predictors for New Observations
New Obs         x
1            50.0
```

Las bandas de confianza y bandas de predicción generadas por *MINITAB* para los datos de calificaciones en cálculo se muestran en la figura 8.15. Observe que en general las bandas de confianza son más angostas que las bandas de predicción para todo valor de puntos x de examen de aprovechamiento. Es seguro que usted esperaría predicciones de que un valor individual fuera mucho más variable que las estimaciones del valor promedio. También observe que las bandas parecen hacerse más anchas a medida que el valor de x_0 se aleja de la media $\bar{x}$. Esto es porque los errores estándar empleados en los intervalos de confianza y predicción contienen el término $(x_0 - \bar{x})^2$, que se hace más grande cuando los dos valores divergen. En la práctica, esto significa que estimación y predicción son más precisos cuando x_0 está cerca del centro del rango de los valores de x. Se pueden localizar los intervalos de confianza y predicción calculados cuando $x = 50$ en la figura 8.15.

FIGURA 8.15
Intervalos de confianza y predicción para los datos de la tabla 8.1



8.7 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

Procedimientos de regresión lineal: *Microsoft Excel*

Ya se utilizaron algunos de los procedimientos de regresión lineal disponibles en *Microsoft Excel* para obtener una gráfica de dispersión de los datos y la recta de regresión de mínimos cuadrados y para calcular el coeficiente de correlación r para un conjunto de datos bivariado. Ahora que ha estudiado las técnicas de prueba y estimación para un análisis de regresión lineal simple, más opciones de *MS Excel* están disponibles para usted.

EJEMPLO 8.9 Consulte la tabla 8.1, en la cual se estudió la relación entre x = puntuación en el examen de aprovechamiento en matemáticas y y = calificación final en cálculo.

Estudiante	Puntuación en el examen de aprovechamiento en matemáticas (x)	Calificación final en cálculo (y)
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

Introduzca los valores para x y y en las columnas A y B de una hoja de cálculo de *Excel*.

1. Use **Data ► Data Analysis ► Regression** para generar el cuadro de diálogo que se muestra en la figura 8.16a. Resalte o teclee los rangos de celdas para los valores de x y y, y seleccione "Labels" si es necesario.
2. Si hace clic en "Confidence Level", *Excel* calculará intervalos de confianza para las estimaciones de regresión, a y b. Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK**.
3. La salida aparecerá en la ubicación de celda seleccionada, y deberá ajustarse usando **Format ► AutoFit Column Width** en la pestaña **Home** en el grupo **Cells** mientras todavía está resaltado. Puede disminuir la precisión decimal si lo desea, usando ←.0 .00 en la pestaña **Home** en el grupo **Number** (véase la figura 8.16b).
4. La salida que se muestra en la figura 8.16b también puede encontrarse en las figuras 8.6b y 8.7b, encontrándose su interpretación en las secciones 8.2 y 8.3 del texto.

FIGURA 8.16

(a)

(b)

D	E	F	G	H	I	J
SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.8398					
R Square	0.7052					
Adjusted R Square	0.6684					
Standard Error	8.7036					
Observations	10					
ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	1	1449.974131	1449.974131	19.141	0.002	
Residual	8	606.025869	75.75323363			
Total	9	2056				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	40.784	8.507	4.794	0.001	21.167	60.401
x	0.766	0.175	4.375	0.002	0.362	1.169

NOTA: *MS Excel* no proporciona opciones para la estimación y predicción ni para la prueba de correlación significativa de la sección 8.5.

Procedimientos de regresión lineal: *MINITAB*

Ya conocemos algunos de los procedimientos de regresión lineal disponibles en *MINITAB*, para obtener una gráfica de la recta de regresión de mínimos cuadrados de mejor ajuste, así como para calcular el coeficiente de correlación r para un conjunto de datos biva-

riados. Ahora que ya hemos estudiado las técnicas de prueba y estimación para un análisis de regresión lineal sencillo, existen más opciones *MINITAB* para usted.

EJEMPLO 8.10

Consulte la tabla 8.1, en la cual se estudió la relación entre x = puntuación en el examen de aprovechamiento en matemáticas y y = calificación final en cálculo.

Estudiante	Puntuación en el examen de aprovechamiento en matemáticas (x)	Calificación final en cálculo (y)
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

Introduzca los valores para x y y en las primeras dos columnas de una hoja de cálculo de *MINITAB*.

1. Las herramientas principales para el análisis de regresión lineal se generan usando **Stat ▶ Regression ▶ Regression**. (Se usará esta misma secuencia de comandos en el capítulo 9 cuando se estudie el *análisis de regresión múltiple*.) El cuadro de diálogo para el comando Regression se muestra en la figura 8.17a.
2. Seleccione **y** para la variable "Response" y **x** para la variable "Predictor". Ahora puede generar algunas gráficas de residuales para comprobar la validez de sus suposiciones de regresión antes de usar el modelo para estimación o predicción. Elija **Graphs** para desplegar el cuadro de diálogo que se muestra en la figura 8.17b. Hemos usado gráficas de residuales **Regular**, seleccionando los cuadros para "Normal plot of residuals" y "Residuals versus fits". Haga clic en **OK** para regresar al cuadro de diálogo principal.

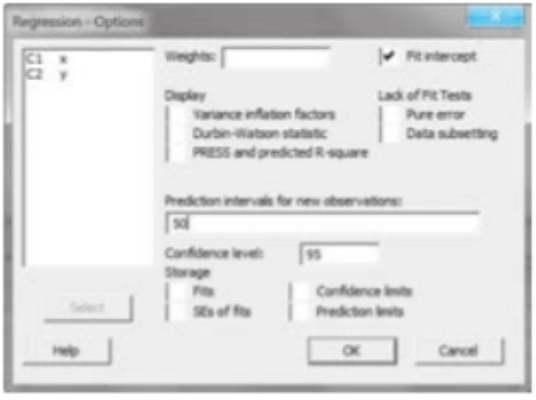
FIGURA 8.17

3. Si ahora elige **Options**, puede obtener intervalos de confianza y predicción para cualquiera de estos casos:
   - Un solo valor de x (introducido en el cuadro marcado "Prediction intervals for new observations").
   - Varios valores de x almacenados en una columna (digamos, C3) de la hoja de cálculo.

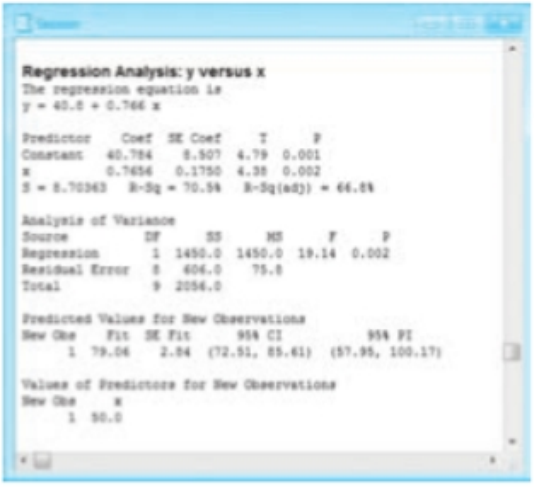
4. Introduzca el valor **50** como en la figura 8.18a para igualar la salida dada en la figura 8.14. Cuando haga clic en **OK** dos veces, se genera la salida de regresión como se muestra en la figura 8.18b. Las dos gráficas de diagnóstico aparecerán en ventanas gráficas separadas.

FIGURA 8.18

(a)



(b)



5. Si lo desea, ahora puede graficar los puntos de datos, la recta de regresión y los límites de confianza y de predicción superiores e inferiores (véase la figura 8.15) usando **Stat ▶ Regression ▶ Fitted Line Plot**. Seleccione y y x para las variables de respuesta y predictora y haga clic en "Display confidence interval" y "Display prediction interval" en el cuadro de diálogo **Options**. Asegúrese de que está seleccionado **Linear** como el "Type of Regression Model", de modo que obtenga un ajuste lineal para los datos.
6. Se utiliza el comando **Stat ▶ Basic Statistics ▶ Correlation** para obtener el valor del coeficiente de correlación r. Asegúrese de que está seleccionado el cuadro marcado "Display p-values". La salida para este comando (usando los datos de prueba/calificación) se muestra en la figura 8.19. Note que el valor p para la prueba de H_0: $\rho = 0$ es idéntico al valor p para la prueba de H_0: $\beta = 0$, ¡porque las pruebas son exactamente equivalentes!

FIGURA 8.19

EJERCICIOS

TÉCNICAS BÁSICAS

8.1 Grafique la recta correspondiente a la ecuación $y = 2x + 1$ al graficar los puntos correspondientes a $x = 0$, 1 y 2. Dé el punto de intersección con el eje y y la pendiente para la recta.

8.2 Dé la ecuación y la gráfica para una recta con intersección con el eje y igual a 3 y pendiente igual a -1.

8.3 ¿Cuál es la diferencia entre modelos matemáticos deterministas y probabilísticos?

8.4 Se le dan cinco puntos con estas coordenadas:

x	−2	−1	0	1	2
y	1	1	3	5	5

a. Use el método de entrada de datos en su calculadora científica o graficadora para introducir las $n = 5$ observaciones. Encuentre las sumas de cuadrados y productos S_{xx}, S_{xy} y S_{yy}.

b. Encuentre la recta de mínimos cuadrados para los datos.

c. Grafique los cinco puntos y grafique la recta de la parte **b**. ¿La recta parece proporcionar un buen ajuste para los puntos de datos?

d. Construya la tabla ANOVA para la regresión lineal.

x	1	2	3	4	5	6
y	5.6	4.6	4.5	3.7	3.2	2.7

8.5 ¿En qué forma el coeficiente de correlación mide la fuerza de la relación lineal entre dos variables y y x?

8.6 Describa la significancia del signo algebraico y la magnitud de r.

8.7 ¿Qué valor toma r si todos los puntos de datos caen en la misma recta en estos casos?

a. La recta tiene pendiente positiva.

b. La recta tiene pendiente negativa.

8.8 Nos dan estos datos:

x	1	2	3	4	5	6
y	7	5	5	3	2	0

a. Grafique los seis puntos en papel para graficar.

b. Calcule el coeficiente muestral de correlación r e interprete.

c. ¿En qué porcentaje se redujo la suma de cuadrados de desviaciones al usar el pronosticador de mínimos cuadrados $\hat{y} = a + bx$, en lugar de $\bar{y}$ como pronosticador de y?

APLICACIONES

8.9 Profesor Asimov El profesor Isaac Asimov fue uno de los escritores más prolíficos de todos los tiempos. Antes de su muerte, escribió casi 500 libros durante una carrera de 40 años. De hecho, conforme su carrera avanzaba fue incluso más productivo en términos del número de libros escritos en un periodo determinado.[1] Los datos siguientes dan el tiempo, en meses, necesario para escribir sus libros en incrementos de 100:

Número de libros, x	100	200	300	400	490
Tiempo en meses, y	237	350	419	465	507

a. Suponga que el número de libros x y el tiempo en meses y están relacionados linealmente. Encuentre la recta de mínimos cuadrados que relacione y con x.

b. Grafique el tiempo como función del número de libros escritos usando una gráfica de dispersión y grafique la recta de mínimos cuadrados en el mismo papel. ¿La parece que la recta da un buen ajuste a los puntos de datos?

c. Construya una tabla ANOVA para la regresión lineal.

Escuela	1	2	3	4	5	6	7	8
API	745	808	798	791	854	688	801	751
EL	71	18	24	50	17	71	11	57

DATOS EX1215

8.10 Entrevistas de prueba De dos técnicas existentes para evaluación de personal, la primera requiere una entrevista de prueba de dos horas mientras que la segunda se puede completar en menos de una hora. Las puntuaciones para cada una de las 15 personas que tomaron ambas pruebas se dan en la tabla siguiente.

Solicitante	Prueba 1 (x)	Prueba 2 (y)
1	75	38
2	89	56
3	60	35
4	71	45
5	92	59
6	105	70
7	55	31
8	87	52
9	73	48
10	77	41
11	84	51
12	91	58
13	75	45
14	82	49
15	76	47

a. Construya una gráfica de dispersión para los datos. ¿Le parece razonable la suposición de linealidad?

b. Encuentre la recta de mínimos cuadrados para los datos.

c. Use la recta de regresión para predecir la puntuación en la segunda prueba para un solicitante que obtuvo 85 puntos en la prueba 1.

DATOS EX1217

8.11 Distancia entre brazos extendidos y estatura Leonardo da Vinci (1452-1519) dibujó la figura de un hombre, indicando que la distancia entre los brazos extendidos de una persona (midiendo por la espalda con los brazos extendidos para formar una "T") es casi igual a la estatura de una persona. Para probar lo dicho por él, medimos ocho personas con los siguientes resultados:

Persona	1	2	3	4
Distancia entre los brazos extendidos (pulgadas)	68	62.25	65	69.5
Estatura	69	62	65	70

Persona	5	6	7	8
Distancia entre los brazos extendidos (pulgadas)	68	69	62	60.25
Estatura	67	67	63	62

a. Trace la gráfica de dispersión para distancia entre los brazos extendidos y estatura. Use la misma escala en los ejes horizontal y vertical. Describa la relación entre las dos variables.

b. Si Da Vinci estaba en lo correcto y la distancia entre los brazos extendidos de una persona es casi igual a la estatura de esa persona, ¿cuál debe ser la pendiente de la recta de regresión?

c. Calcule la recta de regresión para predecir la estatura con base en la distancia entre los brazos extendidos de una persona. ¿El valor de la pendiente b confirma sus conclusiones de la parte **b**?

d. Si una persona tiene una distancia de 62 pulgadas entre los brazos extendidos, ¿cuál sería el pronóstico de usted respecto a la estatura de la persona?

8.12 Langostas La tabla siguiente da los números de lapas *Octolasmis tridens* y *O. lowei* en cada una de 10 langostas.[2] ¿Le parece que las lapas compiten por espacio en la superficie de una langosta?

DATOS EX1253

Número de campo de langosta	*O. tridens*	*O. lowei*
A061	645	6
A062	320	23
A066	401	40
A070	364	9
A067	327	24
A069	73	5
A064	20	86
A068	221	0
A065	3	109
A063	5	350

a. Si compiten, ¿se espera que el número x de lapas *O. tridens* y el número y de *O. lowei* estén correlacionados positiva o negativamente? Explique.

b. Si desea probar la teoría de que los dos tipos de lapas compiten por espacio al realizar una prueba de la hipótesis nula "el coeficiente de correlación poblacional r es igual a 0", ¿cuál es su hipótesis alternativa?

c. Realice la prueba de la parte **b** y exprese sus conclusiones.

DATOS EX1260

8.13 Estadísticas de béisbol ¿El promedio de bateo de un equipo depende en alguna forma del número de cuadrangulares conectados por el equipo? Los datos de la tabla siguiente muestran el número de cuadrangulares del equipo y el promedio general de bateo del equipo para ocho equipos de ligas mayores seleccionados para la temporada de 2010.[3]

Equipo	Cuadrangulares totales	Promedio de bateo del equipo
Atlanta Braves	139	.258
Baltimore Orioles	133	.259
Boston Red Sox	211	.268
Chicago White Sox	177	.268
Houston Astros	108	.247
LA Dodgers	120	.252
Philadelphia Phillies	166	.260
Seattle Mariners	101	.236

Fuente: ESPN.com

a. Grafique los puntos usando una gráfica de dispersión. ¿Le parece que hay alguna relación entre el total de cuadrangulares y el promedio de bateo del equipo?

b. ¿Hay alguna correlación positiva significativa entre el total de cuadrangulares y el promedio de bateo del equipo? Pruebe al nivel de significancia de 5%.

c. ¿Piensa usted que la relación entre estas dos variables sería diferente si hubiéramos visto todo el conjunto de franquicias de las ligas mayores?

DATOS EX1245

8.14 Drew Brees El número de pases completos y el número total de yardas ganadas por pase para Drew Brees, mariscal de campo de los Santos de Nueva Orleans, se registraron en los 16 juegos regulares de la temporada de futbol de 2010.[4] En la semana 10 no se jugó y no hay datos reportados.

Semana	Pases completos	Yardas totales
1	27	237
2	28	254
3	30	365
4	33	275
5	24	279
6	21	263
7	37	356
8	34	305
9	27	253
10	–	–
11	29	382
12	23	352
13	24	313
14	25	221
15	29	267
16	25	302
17	22	196

a. ¿Cuál es la recta de mínimos cuadrados que relaciona el total de yardas ganadas por pase con el número de pases completos para Drew Brees?

b. ¿Qué proporción de la variación total está explicada por la regresión del total de yardas ganadas por pase (y) en el número de pases completos (x)?

c. Si las hay, examine las gráficas de diagnóstico para comprobar la validez de las suposiciones de regresión.

Fuentes

1. Stellan Ohlsson, "The Learning Curve for Writing Books: Evidence from Professor Asimov," *Psychological Science* 3, no. 6 (1992):380–382.
2. W. B. Jeffries, H. K. Voris y C. M. Yang, "Diversity and Distribution of the Pedunculate Barnacles *Octolasmis* Gray, 1825 Epizoic on the Scyllarid Lobster, *Thenus orientalis* (Lund, 1793)," *Crustaceana* 46, no. 3 (1984).
3. http://espn.go.com/mlb/stats/team/_/stat/batting/year/2010/seasontype/2, 5 de noviembre de 2010.
4. "Drew Brees #9 QB, "http://sports.espn.go.com/nfl/players/gamelog?playerId=2580, 22 de marzo de 2011.

9

Regresión múltiple

© Will & Deni McIntyre/CORBIS

OBJETIVOS GENERALES

En este capítulo, extendemos los conceptos de regresión y correlación lineales a una situación donde el valor promedio de una variable aleatoria y está relacionada con varias variables independientes, $x_1, x_2, \ldots, x_k$, en modelos que son más flexibles que el modelo de recta del capítulo 8. Con el análisis de regresión múltiple, podemos usar la información proporcionada por las variables independientes para ajustar varios tipos de modelos a los datos muestrales, para evaluar la utilidad de estos modelos y finalmente para estimar el valor promedio de y o predecir el valor real de y para valores dados de $x_1, x_2, \ldots, x_k$.

ÍNDICE DEL CAPÍTULO

- Análisis de varianza para regresión múltiple
- El análisis de varianza de la prueba F
- El coeficiente de determinación R^2
- El método de mínimos cuadrados
- El modelo y suposiciones lineales generales
- Estimación y predicción usando el modelo de regresión
- Gráficas residuales
- Prueba de los coeficientes de regresión parcial
- R^2 ajustada
- Sumas secuenciales de cuadrados

La **regresión lineal múltiple** es una extensión de la regresión lineal simple para tomar en cuenta más de una variable independiente. Esto es, en lugar de usar sólo una variable independiente x para explicar la variación en y, se usan simultáneamente diversas variables independientes (o elementos de predicción). Con el uso de más de una variable independiente se debe hacer un mejor trabajo para explicar la variación en y y en consecuencia hacer predicciones más precisas.

Por ejemplo, las ventas regionales y del producto de una compañía podrían estar relacionadas con tres factores:

- x_1: la cantidad gastada en publicidad en televisión
- x_2: la cantidad gastada en publicidad en periódicos
- x_3: el número de vendedores asignados a la región

Un investigador recolectaría datos para medir las variables y, x_1, x_2 y x_3, y luego usaría estos datos muestrales para construir una ecuación de predicción que relacionara y con las tres variables predictoras. Desde luego que surgen varias preguntas, al igual que con la regresión lineal simple:

- ¿Qué tan bien se ajusta el modelo?
- ¿Qué tan fuerte es la relación entre y y las variables predictoras?
- ¿Se han violado suposiciones importantes?
- ¿Qué tan buenas son las estimaciones y predicciones?

Para contestar estas preguntas se pueden usar métodos de **análisis de regresión múltiple**, que casi siempre se hacen con un programa de cómputo. Este capítulo contiene una breve introducción al análisis de regresión múltiple y a la difícil tarea de construcción de modelos, es decir, elegir el modelo correcto para una aplicación práctica.

9.1 DESARROLLO DEL MODELO DE REGRESIÓN MÚLTIPLE

El **modelo lineal general** para un análisis de regresión múltiple describe una respuesta particular y usando el modelo que damos a continuación.

MODELO LINEAL GENERAL Y SUPOSICIONES

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon$$

donde

- y es la **variable de respuesta** que se desea predecir.
- β_0, β_1, β_2, ..., β_k son constantes desconocidas.
- $x_1, x_2, \ldots, x_k$ son **variables predictoras** independientes que se miden sin error.
- ϵ es el error aleatorio, que permite que cada respuesta se desvíe del valor promedio de y en una cantidad ϵ. Se debe suponer que los valores de ϵ: 1) son independientes; 2) tienen una media de 0 y una varianza común σ^2 para cualquier conjunto $x_1, x_2, \ldots, x_k$ y 3) están normalmente distribuidos.

Cuando se cumplen estas suposiciones acerca de ϵ, el valor *promedio* de y para un conjunto dado de valores $x_1, x_2, \ldots, x_k$ es igual a la parte *determinista* del modelo:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

Observe que el modelo de regresión simple y las suposiciones son *muy semejantes* al modelo y las suposiciones empleados para la regresión lineal. Es probable que no lo sorprenda el que los procedimientos de prueba y estimación también sean extensiones de los empleados en el capítulo 8.

Los modelos de regresión múltiple son muy flexibles y pueden tomar muchas formas, dependiendo de la manera en que las variables independientes $x_1, x_2, \ldots, x_k$ se introduzcan en el modelo. Empezamos con un simple modelo de regresión múltiple, explicando los conceptos y procedimientos básicos con un ejemplo. A medida que nos familiaricemos con los procedimientos de regresión múltiple, aumentamos la complejidad de los ejemplos y veremos que los mismos procedimientos se pueden usar para modelos de formas diferentes, dependiendo de la aplicación particular.

EJEMPLO 9.1

Suponga que se desea relacionar una variable aleatoria y con dos variables independientes x_1 y x_2. El modelo de regresión múltiple es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

con el valor medio de y dado como

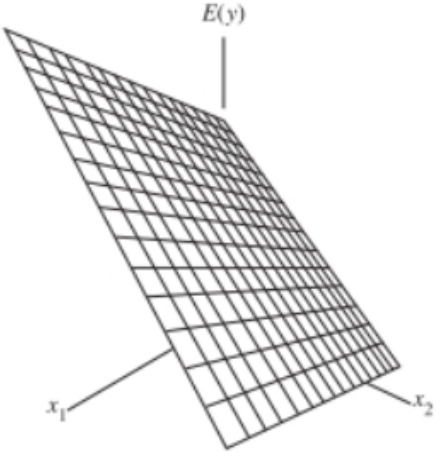
$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Esta ecuación es una extensión en tres dimensiones de la **recta de medias** del capítulo 8 y traza un **plano** en el espacio tridimensional (véase la figura 9.1). La constante β_0 se denomina **punto de cruce**, que es el valor promedio de y cuando x_1 y x_2 son 0 ambas. Los coeficientes β_1 y β_2 se denominan **pendientes parciales** o **coeficientes de regresión parciales**. La pendiente parcial β_i (para $i = 1$ o 2) mide el cambio en y para un cambio unitario en x_i cuando *todas las otras variables independientes se mantienen constantes.* El valor del coeficiente de regresión parcial, por ejemplo β_1, con x_1 y x_2 en el modelo generalmente *no* es igual que la pendiente cuando se ajuste una recta sólo con x_1. Estos coeficientes son las constantes desconocidas, que deben ser estimadas usando datos muestrales para obtener la ecuación de predicción.

TIP MI CONSEJO

En lugar de x y y graficadas en un espacio en dos dimensiones, y y $x_1, x_2, \ldots, x_k$ tienen que graficarse en $(k + 1)$ dimensiones

FIGURA 9.1
Plano de medias para el ejemplo 9.1



Significado de los coeficientes de regresión

Un análisis de regresión múltiple comprende procedimientos de estimación, prueba y diagnóstico diseñados para ajustar el modelo de regresión múltiple

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

a un conjunto de datos. Debido a la complejidad de los cálculos involucrados, estos procedimientos casi siempre se ponen en práctica con un programa de regresión de alguno de los

varios paquetes de software. Todos dan resultados similares en formas ligeramente diferentes. Seguimos con patrones básicos puestos en regresión lineal simple, empezando con un resumen de los procedimientos generales e ilustrado con un ejemplo.

Método de mínimos cuadrados

La ecuación de predicción

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k$$

es la recta que reduce la SSE, la suma de cuadrados de las desviaciones de los valores y observados en los valores pronosticados $\hat{y}$. Estos valores se calculan usando un programa de regresión.

EJEMPLO 9.2

¿En qué forma los vendedores de bienes raíces determinan el precio de venta para un condominio recién inscrito en lista? La base de datos de una computadora en una pequeña comunidad contiene el precio de venta de lista y (en miles de dólares), la cantidad de área de vivienda x_1 (en cientos de pies cuadrados), así como los números de pisos x_2, recámaras x_3 y baños x_4, para $n = 15$ condominios seleccionados al azar actualmente en el mercado. Los datos se muestran en la tabla 9.1.

TABLA 9.1 **Datos sobre 15 condominios**

Observación	Precio de lista, y	Área de vivienda, x_1	Pisos, x_2	Recámaras, x_3	Baños, x_4
1	169.0	6	1	2	1
2	218.5	10	1	2	2
3	216.5	10	1	3	2
4	225.0	11	1	3	2
5	229.9	13	1	3	1.7
6	235.0	13	2	3	2.5
7	239.9	13	1	3	2
8	247.9	17	2	3	2.5
9	260.0	19	2	3	2
10	269.9	18	1	3	2
11	234.9	13	1	4	2
12	255.0	18	1	4	2
13	269.9	17	2	4	3
14	294.5	20	2	4	3
15	309.9	21	2	4	3

El modelo de regresión múltiple es

$$E(y) = \beta_0 = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4$$

que se ajusta usando los paquetes de software *MINITAB* o *Microsoft Excel*. En la sección 9.6 se encuentran instrucciones para generar esta salida. La primera parte de la salida de regresión de *MINITAB* se muestra en la figura 9.2a. Se encontrará la ecuación de regresión ajustada en los primeros dos renglones de la salida impresa:

$$\hat{y} = 119 + 6.27x_1 - 16.2x_2 - 2.67x_3 + 30.3x_4$$

Los coeficientes de regresión parcial se muestran con ligeramente más precisión en la segunda sección de la salida impresa de *MINITAB*, una salida similar generada por *MS Excel* se muestra en la figura 9.2b. Las columnas son una lista del nombre dado a cada variable independiente de pronóstico, su coeficiente de regresión estimado, su error estándar y los valores t y p que se usan para probar su significancia *en presencia de todas las otras variables predictoras*. En una sección más adelante explicamos estas pruebas con más detalle.

FIGURA 9.2a

Una porción de la salida impresa *MINITAB* para el ejemplo 9.2

Análisis de regresión: precio de lista contra pies cuadrados, número de pisos, recámaras, baños

```
The regression equation is
List Price = 119 + 6.27 Square Feet - 16.2 Number of Floors
           - 2.67 Bedrooms + 30.3 Baths

Predictor              Coef     SE Coef          T        P
Constant            118.763       9.207      12.90    0.000
Square Feet          6.2698      0.7252       8.65    0.000
Number of Floors    -16.203       6.212      -2.61    0.026
Bedrooms             -2.673       4.494      -0.59    0.565
Baths                30.271       6.849       4.42    0.001
```

FIGURA 9.2b

Una porción de la salida impresa de *MS Excel* para el ejemplo 9.2

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Punto de cruce	118.763	9.207	12.899	0.000	98.248	139.279
Pies cuadrados	6.270	0.725	8.645	0.000	4.654	7.886
Número de pisos	-16.203	6.212	-2.608	0.026	-30.045	-2.362
Recámaras	-2.673	4.494	-0.595	0.565	-12.686	7.340
Baños	30.271	6.849	4.420	0.001	15.011	45.530

Análisis de varianza para regresión múltiple

El análisis de varianza divide la variación total en la variable de respuesta y,

$$\text{SS Total} = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}$$

en dos partes:

- La SSR (suma de cuadrados para regresión) mide la cantidad de variación explicada usando la ecuación de regresión.
- La SSE (suma de cuadrados para error) mide la variación residual en los datos que no está explicada por las variables independientes.

de modo que

$$\text{SS total} = \text{SSR} + \text{SSE}$$

Los **grados de libertad** para estas sumas de cuadrados se encuentran usando el siguiente argumento. Hay $(n - 1)$ grados de libertad en total. Estimar la recta de regresión requiere estimar k coeficientes desconocidos: $\beta_1, \beta_2, \ldots, \beta_k$; la constante b_0 (la cual estima β_0) es una función de $\bar{y}$ y de las otras estimaciones. En consecuencia, hay k grados de libertad de regresión, dejando $(n - 1) - k$ grados de libertad para error. Al igual que en capítulos anteriores, las medias cuadráticas se calculan como $\text{MS} = \text{SS}/gl$.

La tabla ANOVA para los datos de bienes raíces de la tabla 9.1 se muestra en la segunda porción de la salida impresa *MINITAB* en la figura 9.3a y en la sección inferior de la salida impresa de *Excel* en la figura 9.3b. Hay $n = 15$ observaciones y $k = 4$ variables predictoras independientes. Se puede verificar que el total de grados de libertad, $(n - 1) = 14$, se divide en $k = 4$ para regresión y $(n - k - 1) = 10$ para error.

FIGURA 9.3a

Una porción de la salida impresa *MINITAB* para el ejemplo 9.2

```
S = 6.84930       R-Sq = 97.1%     R-Sq(adj) = 96.0%

Analysis of Variance
Source            DF          SS          MS         F        P
Regression         4     15913.0      3978.3     84.80    0.000
Residual Error    10       469.1        46.9
Total             14     16382.2

Source               DF      Seq SS
Square Feet           1     14829.3
Number of Floors      1         0.9
Bedrooms              1       166.4
Baths                 1       916.5
```

FIGURA 9.3b

Una porción de la salida impresa de *MS Excel* para el ejemplo 9.2

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.986
R Square	0.971
Adjusted R Square	0.960
Standard Error	6.849
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	15913.048	3978.262	84.801	0.000
Residual	10	469.129	46.913		
Total	14	16382.177			

La mejor estimación de la variación aleatoria σ^2 en el experimento, es decir la variación que no es explicada por las variables predictoras, como de costumbre está dada por

$$s^2 = \text{MSE} = \frac{\text{SSE}}{n - k - 1} = 46.913$$

de la tabla ANOVA. El primer renglón de la figura 9.3a y el cuarto renglón en la figura 9.3b también muestran $s = \sqrt{s^2} = 6.849$. La computadora usa estos valores internamente para producir estadísticos de prueba, intervalos de confianza e intervalos de predicción, que estudiaremos en secciones subsiguientes.

La última sección de la figura 9.3a muestra una descomposición de SSR = 15 913.0 en que la contribución condicional de cada variable de predicción *dadas las variables ya introducidas en el modelo* se muestra para el orden de entrada que se especifique en el programa de regresión. Para el ejemplo de bienes raíces, el programa *MINITAB* introdujo las variables en este orden: pies cuadrados, seguido de números de pisos, recámaras y baños. Cada una de estas **sumas de cuadrados secuenciales** o condicionales explican uno de los $k = 4$ grados de libertad de regresión. Es interesante observar que la variable de predicción x_1 por sí sola da cuenta de 14 829.3/15 913.0 = .932 o 93.2% de la variación total explicada por el modelo de regresión, pero, si cambia el orden de entrada, otra variable puede explicar la mayor parte de la suma de cuadrados de regresión.

Prueba de la utilidad del modelo de regresión

Ya se probó si y y x estaban linealmente relacionadas al probar H_0: $\beta = 0$ con una prueba t o una prueba F equivalente. En regresión múltiple, hay más de una *pendiente parcial*, que son los *coeficientes de regresión parcial.* Las pruebas t y F ya no son equivalentes.

9.2 EL COEFICIENTE DE DETERMINACIÓN MÚLTIPLE, PRUEBA *F* GLOBAL

TIP MI CONSEJO

Las salidas impresas *MINITAB* informan de R^2 como un porcentaje más que una proporción

¿Qué tan bien se ajusta el modelo de regresión? La salida impresa de regresión da una medida estadística de la fuerza del modelo en el **coeficiente de determinación**, $\boldsymbol{R^2}$; es decir, la proporción de la variación total que es explicada por la regresión de y en $x_1, x_2, \ldots, x_k$, definida como

$$R^2 = \frac{\text{SSR}}{\text{SS Total}} = \frac{15\,913.0}{16\,382.2} = .971 \quad \text{o} \quad 97.1\%$$

El coeficiente de determinación se denomina a veces $\boldsymbol{R^2}$ **múltiple** y se encuentra en el primer renglón de la figura 9.3a, marcado "R-Sq" y en la segunda línea de la figura 9.3b, etiquetada "R Square". En consecuencia, para el ejemplo de bienes raíces, 97.1% de la variación total ha sido explicado por el modelo de regresión. El modelo se ajusta muy bien.

TIP MI CONSEJO

R^2 es la multivariada equivalente de r^2, empleada en regresión lineal

Puede ser útil saber que el valor del estadístico F está relacionado con R^2 por la fórmula

$$F = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

de manera que R^2 es grande, F es grande y viceversa.

Análisis de varianza de la prueba *F*

La ecuación de regresión que usa información dada por las variables predictoras $x_1, x_2, ..., x_k$ ¿es considerablemente mejor que la predictora simple $\bar{y}$ que no se apoya en ninguno de los valores de x? Esta pregunta se contesta usando una prueba F general con las hipótesis:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

contra

TIP MI CONSEJO

La prueba *F* general (para la significancia del modelo) en regresión múltiple es de una cola

H_a: Al menos una de $\beta_1, \beta_2, \ldots, \beta_k$ no es 0

El estadístico de prueba se encuentra en la tabla ANOVA (figura 9.3) como

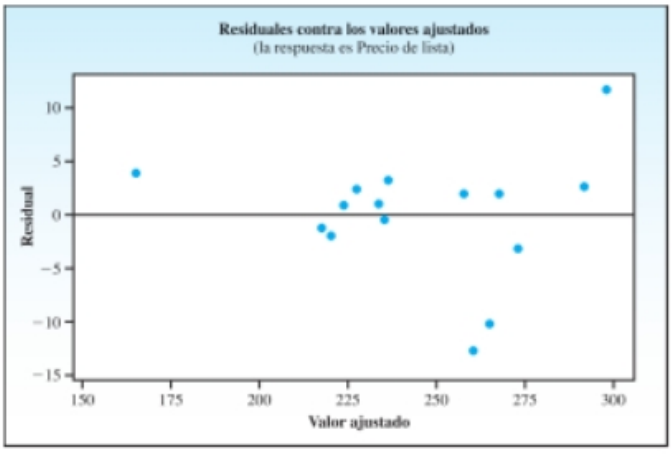
$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{3978.3}{46.9} = 84.80$$

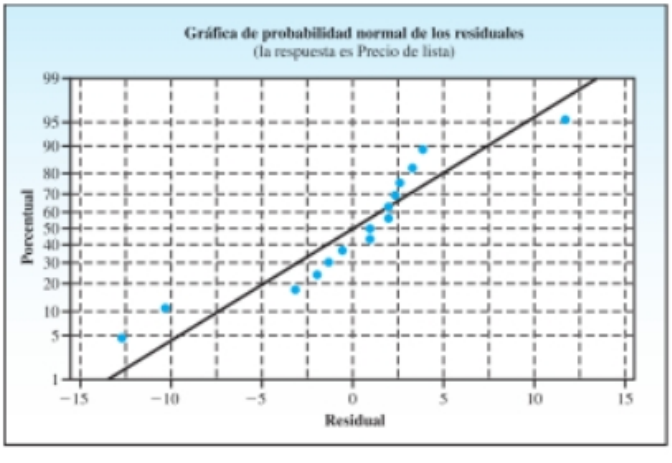
que tiene una distribución F con $gl_1 = k = 4$ y $gl_2 = (n - k - 1) = 10$. Como el valor p exacto, $P = .000$, está dado en la salida impresa, se puede declarar que la regresión es altamente significativa. Esto es, al menos una de las variables predictoras está aportando información significativa para la predicción de la variable de respuesta y.

9.3 ANÁLISIS DE RESIDUALES

Antes de usar el modelo de regresión para su propósito principal, que es estimar y predecir y, deben verse **gráficas residuales** generadas por computadora para asegurarse que sean válidas todas las suposiciones de regresión. La *gráfica normal de probabilidad* y la *gráfica de residuales contra ajuste* se presentan en la figura 9.4 para los datos de bienes raíces. Parece haber tres observaciones que no se ajustan al patrón general. Se pueden ver como resultados atípicos en ambas gráficas. Es probable que estas tres observaciones deban investigarse, pero no dan fuerte evidencia de que las suposiciones se han violado.

FIGURA 9.4
Gráficas de diagnóstico





INFERENCIAS RESPECTO A LOS COEFICIENTES DE REGRESIÓN DE LA POBLACIÓN

9.4

Pruebas de hipótesis

TIP MI CONSEJO

Se puede demostrar que

$$F = \frac{\text{MSR}}{\text{MSE}} = \frac{R^2/k}{(1 - R^2)/(n - k - 1)}$$

Pruebas de la significancia de los coeficientes de regresión parcial

Una vez que hayamos determinado que el modelo es útil para predecir y, debemos explorar la naturaleza de la "utilidad" en más detalle. ¿Todas las variables predictoras agregan información importante para la predicción *en presencia de otras variables predictoras que ya están en el modelo?* Las pruebas t individuales de la primera sección de la salida impresa de regresión están diseñadas para probar las hipótesis

$$H_0: \beta_i = 0 \quad \text{contra} \quad H_a: \beta_i \neq 0$$

para cada uno de los coeficientes de regresión parcial, *dado que las otras variables predictoras ya están en el modelo.* Estas pruebas están basadas en el estadístico t de Student dado por

$$t = \frac{b_i - \beta_i}{\text{SE}(b_i)}$$

que tiene $gl = (n - k - 1)$ grados de libertad. El procedimiento es idéntico al empleado para probar una hipótesis acerca de la pendiente β del modelo de regresión lineal simple.[†]

TIP MI CONSEJO

Pruebe la significancia del coeficiente individual β_i, usando pruebas t

La figura 9.5 muestra las pruebas t y los valores p de la parte superior de la salida impresa de *MINITAB* y la sección inferior de la salida impresa de *MS Excel*. Al examinar los valores p de la última columna se puede ver que todas las variables *excepto* x_3, el número de recámaras, agregan información significativa para predecir y, **aun cuando todas las otras variables independientes ya estén en el modelo**. ¿Podría ser mejor el modelo? Tal vez x_3 sea una variable de predicción innecesaria. Una opción es eliminarla y reajustar el modelo con un nuevo conjunto de datos.

FIGURA 9.5a

Una porción de la salida impresa *MINITAB* para el ejemplo 9.2

Predictor	Coef	SE Coef	T	P
Constant	118.763	9.207	12.90	0.000
Square Feet	6.2698	0.7252	8.65	0.000
Number of Floors	-16.203	6.212	-2.61	0.026
Bedrooms	-2.673	4.494	-0.59	0.565
Baths	30.271	6.849	4.42	0.001

FIGURA 9.5b

Una porción de la salida impresa de *MS Excel* para el ejemplo 9.2

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	118.763	9.207	12.899	0.000
Square Feet	6.270	0.725	8.645	0.000
Number of Floors	-16.203	6.212	-2.608	0.026
Bedrooms	-2.673	4.494	-0.595	0.565
Baths	30.271	6.849	4.420	0.001

El valor de R^2 ajustado

Observe de la definición de R^2 = SSR/SS Total que su valor nunca puede disminuir con la adición de más variables en el modelo de regresión. En consecuencia, R^2 puede estar artificialmente inflada por la inclusión de más y más variables predictoras.

Una medida alternativa de la fuerza del modelo de regresión se ajusta para grados de libertad con el uso de medias cuadráticas en lugar de sumas de cuadrados:

$$R^2(\text{adj}) = \left(1 - \frac{\text{MSE}}{\text{SS Total}/(n-1)}\right)100\%$$

[†] Algunos paquetes usan el estadístico t que acabamos de describir, mientras que otros usan el estadístico F equivalente ($F = t^2$), puesto que el cuadrado de un estadístico t con v grados de libertad es igual a un estadístico F con 1 gl en el numerador y v grados de libertad en el denominador.

TIP MI CONSEJO

Use R^2(adj) para comparar uno o más modelos posibles

Para los datos de bienes raíces de la figura 9.3,

$$R^2(\text{adj}) = \left(1 - \frac{46.9}{16\,382.2/14}\right)100\% = 96.0\%$$

se encuentra en el primer renglón de la salida impresa de *MINITAB* y en el tercer renglón de la salida impresa de *Excel*. El valor "R-Sq(adj) = 96.0%" o "Adjusted R Square = 0.960" representan el porcentaje de variación en la respuesta y explicada por las variables independientes, corregida para grados de libertad. El valor ajustado de R^2 se usa principalmente para comparar dos o más modelos de regresión que usan números diferentes de variables predictoras independientes.

Intervalos de confianza

TIP MI CONSEJO

Para valores dados de $x_1, x_2, \ldots, x_k$, el intervalo de predicción **siempre** será más ancho que el intervalo de confianza

Finalmente, una vez que se haya determinado que el modelo es efectivo para describir la relación entre y y las variables predictoras $x_1, x_2, \ldots, x_k$, el modelo se puede usar para estos fines:

- Estimar el valor promedio de y, $E(y)$, para valores dados de $x_1, x_2, \ldots, x_k$
- Predecir un valor particular de y para valores dados de $x_1, x_2, \ldots, x_k$

Los valores de $x_1, x_2, \ldots, x_k$ se introducen en la computadora y ésta genera el valor ajustado $\hat{y}$ junto con su error estándar estimado y los intervalos de confianza y predicción. Recuerde que el intervalo de predicción es *siempre más ancho* que el intervalo de confianza.

Veamos qué tan bien funciona nuestra predicción para los datos de bienes raíces, usando otra casa de la base de datos de la computadora, una casa con 1 000 pies cuadrados de superficie de vivienda, un piso, tres recámaras y dos baños, con precio de lista de 221 500 dólares. La salida impresa de la figura 9.6 muestra los intervalos de confianza y predicción para estos valores. El valor real cae dentro de ambos intervalos, lo que indica que el modelo funciona muy bien.

FIGURA 9.6

Intervalos de confianza y predicción para el ejemplo 9.2

```
Predicted Values for New Observations
New Obs      Fit     SE Fit          95% CI              95% PI
1         217.78       3.11   (210.86,  224.70)   (201.02,  234.54)

Values of Predictors for New Observations
New      Square     Number of
Obs       Feet       Floors     Bedrooms      Baths
1         10.0         1.00         3.00       2.00
```

9.5 El MODELO DE REGRESIÓN CUADRÁTICA

En la sección 9.1 explicamos en detalle las diversas partes de la salida impresa de regresión múltiple. Cuando se efectúa un análisis de regresión múltiple, se debe usar un método paso a paso:

1. Obtener el modelo de predicción ajustada.
2. Usar el análisis de varianza de la prueba F y R^2 para determinar qué tan bien se ajusta el modelo a los datos.
3. Verificar las pruebas t para los coeficientes de regresión parcial para ver cuáles están aportando información significativa en presencia de los otros.
4. Si se escoge comparar varios modelos diferentes, use R^2(adj) para comparar su efectividad.
5. Usar gráficas residuales generadas por computadora para ver si hay violación de las suposiciones de regresión.

Una vez que se han tomado todos estos pasos estamos listos para usar el modelo en la estimación y predicción.

Las variables predictoras $x_1, x_2, \ldots, x_k$ empleadas en el modelo lineal general no tienen que representar variables predictoras *diferentes*. Por ejemplo, si se sospecha que una variable in-

dependiente x afecta la respuesta y, pero que la relación es *curvilínea* más que *lineal*, entonces se podría elegir ajustar a un **modelo cuadrático**:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

El modelo cuadrático es un ejemplo de un **modelo de segundo orden** porque contiene un término cuyos exponentes suman 2 (en este caso, x^2).† También es un ejemplo de un **modelo polinomial**, un modelo que toma la forma

$$y = a + bx + cx^2 + dx^3 + \cdots$$

Para ajustar este tipo de modelo usando el programa de regresión múltiple, los valores observados de y, x y x^2 se introducen en la computadora y la salida impresa se puede generar como en la sección 9.1.

TIP MI CONSEJO

Una ecuación cuadrática es $y = a + bx + cx^2$. La gráfica forma una **parábola**

EJEMPLO 9.3

En un estudio de variables que afecta la productividad en el comercio de comestibles al menudeo, W. S. Good usa valor agregado por hora de trabajo para medir la productividad de tiendas de comestibles al menudeo.† Él define "valor agregado" como "el excedente [dinero generado por el negocio] disponible para pagar empleados, mobiliario y enseres, y equipo". Los datos consistentes con la relación entre valor agregado por hora de trabajo y y el tamaño x de una tienda de comestibles descrita en el artículo de Good, se muestran en la tabla 9.2 para 10 tiendas de alimentos ficticias. Elija un modelo para relacionar y con x.

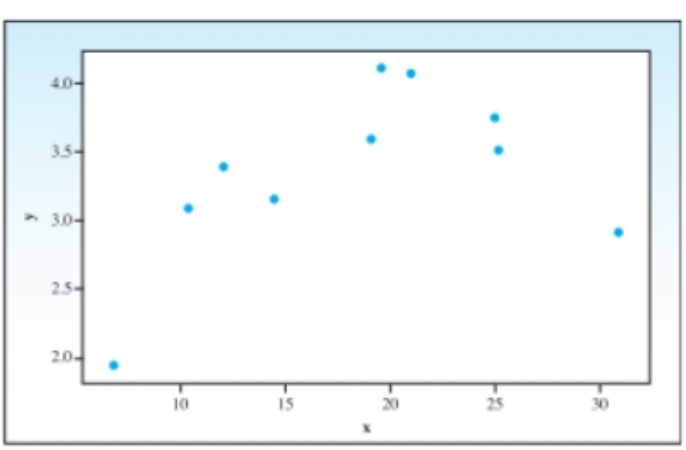
TABLA 9.2 Datos sobre tamaño de tienda y valor agregado

Tienda	Valor agregado por hora de trabajo, y ($)	Tamaño de tienda a (miles de pies cuadrados), x
1	4.08	21.0
2	3.40	12.0
3	3.51	25.2
4	3.09	10.4
5	2.92	30.9
6	1.94	6.8
7	4.11	19.6
8	3.16	14.5
9	3.75	25.0
10	3.60	19.1

Solución Se puede investigar la relación entre y y x al ver la gráfica de los puntos de datos en la figura 9.7. La gráfica sugiere que la productividad, y, aumenta cuando el tamaño de la tienda de comestibles, x, aumenta hasta alcanzar un tamaño óptimo. Arriba de ese tamaño, la productividad tiende a disminuir. La relación parece ser *curvilínea* y un modelo cuadrático,

$$E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$$

FIGURA 9.7
Gráfica del tamaño de la tienda x y valor agregado y para el ejemplo 9.3



† El *orden* de un término está determinado por la suma de los exponentes de variables que conforman ese término. Los términos que contienen x_1 o x_2 son de primer orden; los que contienen x_1^2, x_2^2 o x_1x_2 son de segundo orden.

puede ser apropiado. Recuerde que al elegir usar este modelo, no decimos que la verdadera relación sea cuadrática sino sólo que puede dar estimaciones y predicciones más precisas que, por ejemplo, un modelo lineal.

EJEMPLO 9.4

Consulte los datos sobre productividad y tamaño de una tienda minorista de alimentos del ejemplo 9.3. Se utilizó *MINITAB* para ajustar un modelo cuadrático a los datos y para graficar la curva cuadrática de predicción, junto con los puntos de datos graficados. Analice lo adecuado del modelo ajustado.

Solución De la salida impresa de la figura 9.8 se puede ver que la ecuación de regresión es

$$\hat{y} = -.159 + .392x - .00949x^2$$

La gráfica de esta ecuación cuadrática junto con los puntos de datos se muestra en la figura 9.9.

FIGURA 9.8
Salida impresa *MINITAB* para el ejemplo 9.4

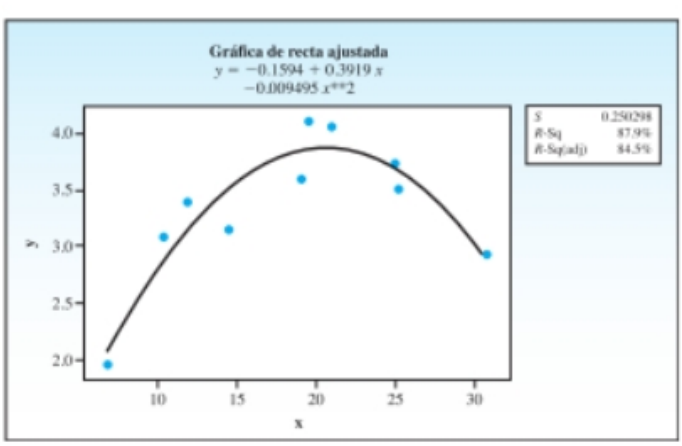
Análisis de regresión: *y* contra *x*, *x*-sq

```
The regression equation is
y = - 0.159 + 0.392 x - 0.00949 x-sq

Predictor         Coef      St Coef          T        P
Constant       -0.1594       0.5006      -0.32    0.760
x              0.39193      0.05801       6.76    0.000
x-sq         -0.009495     0.001535      -6.19    0.000

S = 0.250298       R-Sq = 87.9%      R-Sq(adj) = 84.5%

Analysis of Variance
Source            DF        SS        MS        F        P
Regression         2    3.1989    1.5994    25.53    0.001
Residual Error     7    0.4385    0.0626
Total              9    3.6374

Source       DF      Seq SS
x             1      0.8003
x-sq          1      2.3986
```

TIP MI CONSEJO

Vea la salida impresa y encuentre las leyendas para "Predictor". Esto le dirá cuáles variables se han usado en el modelo

FIGURA 9.9
Recta de regresión cuadrática ajustada para el ejemplo 9.4



Para evaluar lo adecuado del modelo cuadrático, la prueba de

$$H_0: \beta_1 = \beta_2 = 0$$

contra

$$H_a: \text{Ni } \beta_1 \text{ ni } \beta_2 \text{ son } 0$$

se da en la salida impresa como

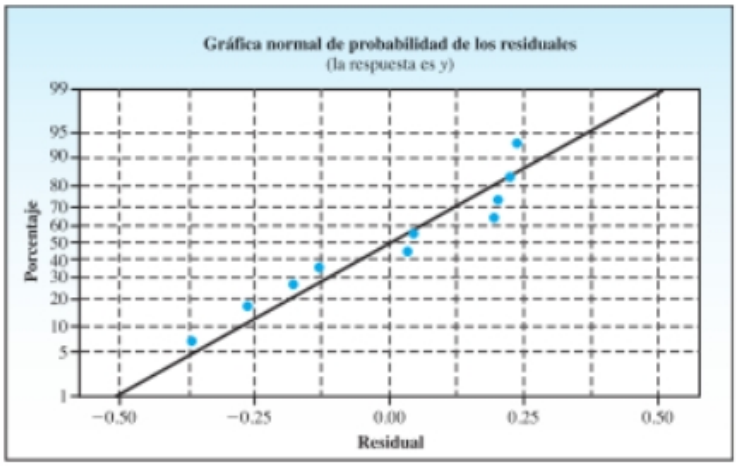
$$F = \frac{\text{MSR}}{\text{MSE}} = 25.53$$

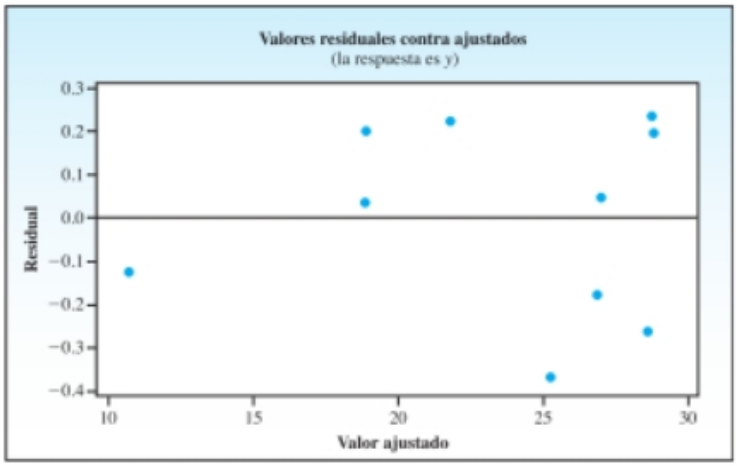
con valor $p = .001$. En consecuencia, el ajuste total del modelo es altamente significativo. La regresión cuadrática explica $R^2 = 87.9\%$ de la variación en y [R^2(adj) = 84.5%].

De las pruebas t para las variables individuales del modelo se puede ver que b_1 y b_2 son altamente significativas, con valores p iguales a .000. Observe de la sección de suma secuencial de cuadrados que la suma de cuadrados para regresión lineal es .8003, con una suma de cuadrados adicional de 2.3986 cuando se agregue el término cuadrático. Es evidente que el modelo de regresión lineal simple es inadecuado para describir los datos.

Una última mirada a las gráficas residuales en la figura 9.10 asegura que las suposiciones de regresión sean válidas. Observe el aspecto relativamente lineal de la gráfica normal y la relativa dispersión de los residuales contra los ajustes. El modelo cuadrático da predicciones precisas para valores de x que se encuentren *dentro del rango de los valores muestreados de x*.

FIGURA 9.10
Gráficas de diagnóstico para el ejemplo 9.4





9.6 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

Procedimientos de regresión múltiple: *Microsoft Excel*

El procedimiento para realizar un análisis de regresión múltiple en *MS Excel* es idéntico al procedimiento de regresión lineal descrito en la sección 8.7, excepto que el rango de las variables x abarcará más de una columna. Podría desear repasar esa sección antes de continuar.

EJEMPLO 9.5

Suponga que una variable de respuesta y se relaciona con cuatro variables predictoras, x_1, x_2, x_3 y x_4, de modo que $k = 4$.

1. Introduzca los valores observados de y y cada una de las $k = 4$ variables predictoras en las primeras $(k + 1)$ columnas de una hoja de cálculo de *Excel*. (NOTA: A fin de que el análisis de regresión múltiple funcione de manera apropiada, debe haber una columna de valores para cada variable predictora independiente x_i en su modelo, y las columnas de las x **deben ser adyacentes entre sí**.)
2. Use **Data ▶ Data Analysis ▶ Regression** para generar el cuadro de diálogo, resaltando o tecleando los rangos de celdas para los valores x_i y y, y seleccione "Labels" de ser necesario.
3. Si hace clic en "Confidence Level", *Excel* calculará intervalos de confianza para las estimaciones de regresión, b_0, b_1, b_2, b_3 y b_4. Introduzca una ubicación de celda para el **Output Range** y haga clic en **OK** para generar la salida de regresión.

   NOTA: *MS Excel* no proporciona opciones para estimación y predicción. Además, las gráficas de diagnóstico que pueden generarse en *Excel* no son las mismas gráficas que se han expuesto en la sección 9.1 y no se comentarán en esta sección.

Procedimientos de regresión múltiple: *MINITAB*

El procedimiento para ejecutar un análisis de regresión múltiple en *MINITAB* es parecido al procedimiento de regresión lineal descrito en la sección 8.7, excepto que el rango de las variables x abarcará más de una columna. Podría desear repasar esa sección antes de continuar.

EJEMPLO 9.6

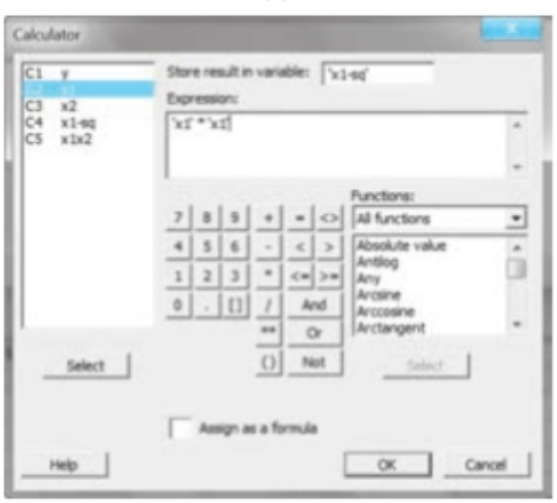
Suponga que una variable de respuesta y se relaciona con cuatro variables predictoras, x_1, x_2, x_3 y x_4, de modo que $k = 4$.

1. Introduzca los valores observados de y y cada una de las $k = 4$ variables predictoras en las primeras $(k + 1)$ columnas de una hoja de cálculo de *MINITAB*. Una vez hecho esto, las principales herramientas inferenciales para el análisis de regresión múltiple se generan usando **Stat ▶ Regression ▶ Regression**. El cuadro de diálogo para el comando **Regression** se muestra en la figura 9.11a.
2. Seleccione **y** para la variable Response y $\mathbf{x_1, x_2, \ldots, x_k}$ para las variables Predictor. Ahora puede elegir generar algunas gráficas residuales para comprobar la validez de sus suposiciones de regresión antes de usar el modelo para estimación o predicción. Elija **Graphs** para mostrar el cuadro de diálogo para gráficas residuales, y elija la gráfica de diagnóstico apropiada.

FIGURA 9.11

(a)

(b)



3. Una vez que haya verificado la pertinencia de su modelo de regresión múltiple, puede elegir **Options** y obtener intervalos de confianza y predicción para cualquiera de estos casos:
   - Un solo conjunto de valores $x_1, x_2, ..., x_k$ (introducidos en el cuadro marcado "Prediction intervals for new observations")
   - Varios conjuntos de valores $x_1, x_2, ..., x_k$ almacenados en k columnas de la hoja de cálculo.

   Cuando haga clic en **OK** dos veces se generará la salida de regresión.
4. La única dificultad en la ejecución del análisis de regresión múltiple usando *MINITAB* podría ser introducir de manera apropiada los datos para su modelo particular. Si el modelo implica términos polinomiales o términos de interacción, el comando **Calc ▶ Calculator** le ayudará. Por ejemplo, suponga que desea ajustar el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_1 x_2$$

   Necesitará introducir los valores observados de y, x_1 y x_2 en las primeras tres columnas de la hoja de cálculo *MINITAB*. Nombre la columna C4 "x1-sq" y nombre C5 "x1x2". Ahora puede usar el cuadro de diálogo de la calculadora que se muestra en la figura 9.11b para generar estas dos columnas. En el cuadro **Expression**, seleccione **x1 * x1** o **x1 ** 2** y almacene los resultados en **C4** (x1-sq). Haga clic en **OK**. Del mismo modo, para obtener los datos para C5, seleccione **x1 * x2** y almacene los resultados en **C5** (x1x2). Haga clic en **OK**. Ahora está listo para ejecutar el análisis de regresión múltiple.
5. Si está ajustando ya sea un modelo cuadrático o un modelo cúbico en una variable x, ahora puede graficar los puntos de datos, la curva de regresión polinomial y los límites de confianza y predicción superior e inferior usando **Stat ▶ Regression ▶ Fitted line Plot**. Seleccione y y x para las variables Response y Predictor, y haga clic en "Display confidence interval" y "Display prediction interval" en el cuadro de diálogo **Options**. Asegúrese de que está seleccionado **Quadratic** o **Cubic** como el "Type of Regression Model", de modo que obtenga el ajuste apropiado para los datos.
6. Recuerde que en el capítulo 8 usó **Stat ▶ Basic Statistics ▶ Correlation** para obtener el valor del coeficiente de correlación r. En el análisis de regresión múltiple, el mismo comando generará una matriz de correlaciones, una por cada par de variables en el conjunto $y, x_1, x_2, ..., x_k$. Asegúrese de que está seleccionado el cuadro marcado "Display p-values". Los valores p proporcionarán información sobre la correlación significativa entre un par particular, en presencia de todas las otras variables en el modelo, y son idénticas a los valores p para las pruebas t individuales de los coeficientes de regresión.

EJERCICIOS

TÉCNICAS BÁSICAS

9.1 Suponga que $E(y)$ está relacionada con dos variables predictoras, x_1 y x_2, por la ecuación

$$E(y) = 3 + x_1 - 2x_2$$

a. Grafique la relación entre $E(y)$ y x_1 cuando $x_2 = 2$. Repita para $x_2 = 1$ y para $x_2 = 0$.

b. ¿Qué relación tienen entre sí las rectas de la parte **a**?

9.2 Suponga que se ajusta el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

a 15 puntos de datos y se encuentra que F es igual a 57.44.

a. ¿Los datos dan suficiente evidencia para indicar que el modelo aporta información para la predicción de y? Pruebe usando un nivel de significancia de 5%.

b. Use el valor de F para calcular R^2. Interprete su valor.

9.3 La salida impresa de computadora para el análisis de regresión múltiple para el ejercicio 9.3 da esta información:

$b_0 = 1.04$ $\quad b_1 = 1.29$ $\quad \text{SE}(b_1) = .42$

$b_2 = 2.72$ $\quad b_3 = .41$ $\quad \text{SE}(b_2) = .65$ $\quad \text{SE}(b_3) = .17$

a. ¿Cuál, si hay alguna, de las variables independientes x_1, x_2 y x_3 aportan información para la predicción de y?

b. Dé la ecuación de predicción de mínimos cuadrados.

c. En la misma hoja de papel, grafique y contra x_1 cuando $x_2 = 1$ y $x_3 = 0$, y cuando $x_2 = 1$ y $x_3 = .5$.

d. ¿Qué relación tienen entre sí las dos rectas?

e. ¿Cuál es la interpretación práctica del parámetro β_1?

9.4 Suponga que se ajusta el modelo

$$E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$$

a 20 puntos de datos y se obtiene la salida impresa *MINITAB* siguiente:

Salida impresa *MINITAB* para el ejercicio 9.4

Análisis de regresión: y contra x, x-sq

```
The regression equation is
y = 10.6 + 4.44 x - 0.648 x-sq

Predictor       Coef    SE Coef       T      P
Constant     10.5638     0.6951   15.20  0.000
x             4.4366     0.5150    8.61  0.000
x-sq         -0.64754    0.07988  -8.11  0.000

S = 1.191    R-Sq = 81.5%    R-Sq(adj) = 79.3%

Analysis of Variance
Source           DF       SS       MS      F      P
Regression        2  106.072   53.036  37.37  0.000
Residual Error   17   24.128    1.419
Total            19  130.200
```

a. ¿Qué tipo de modelo se ha elegido para ajustar los datos?

b. ¿Qué tan bien ajusta los datos el modelo? Explique.

c. ¿Los datos dan suficiente evidencia para indicar que el modelo aporta información para la predicción de y? Use el método del valor p.

9.5 Consulte el ejercicio 9.4.

a. ¿Cuál es la ecuación de predicción?

b. Grafique la ecuación de predicción sobre el intervalo $0 \leq x \leq 6$.

9.6 Consulte el ejercicio 9.4.

a. ¿Cuál es su estimación del valor promedio de y cuando $x = 0$?

b. ¿Los datos dan suficiente evidencia para indicar que el valor promedio de y difiere de 0 cuando $x = 0$?

9.7 Consulte el ejercicio 9.4.

a. Suponga que la relación entre $E(y)$ y x es una recta. ¿Qué se sabría acerca del valor de β_2?

b. ¿Los datos dan suficiente evidencia para indicar curvatura en la relación entre y y x?

APLICACIONES

DATOS EX1310

9.8 Libros de texto universitarios Un editor de libros de texto universitarios realizó un estudio para relacionar la utilidad por texto y, con el costo de ventas x, en un periodo de 6 años cuando su fuerza de ventas (y costos de ventas) estaban creciendo rápidamente. Se recolectaron estos datos de inflación ajustada (en miles de dólares):

Utilidad por texto, y	16.5	22.4	24.9	28.8	31.5	35.8
Costo de ventas por texto, x	5.0	5.6	6.1	6.8	7.4	8.6

Esperando que la utilidad por libro subiera y luego se nivelara, el editor ajustó el modelo $E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$ a los datos.

Salida de *Excel* para el ejercicio 9.8.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.9978
R Square	0.9955
Adjusted R Square	0.9925
Standard Error	0.5944
Observations	6

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	234.995	117.478	332.528	0.000
Residual	3	1.060	0.353		
Total	5	236.015			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-44.192	8.287	-5.333	0.013
x	16.334	2.490	6.560	0.007
x-sq	-0.820	0.182	-4.494	0.021

a. Grafique los puntos de datos. ¿Le parece que el modelo cuadrático es necesario?

b. Encuentre s en la salida impresa. Confirme que

$$s = \sqrt{\frac{\text{SSE}}{n - k - 1}}$$

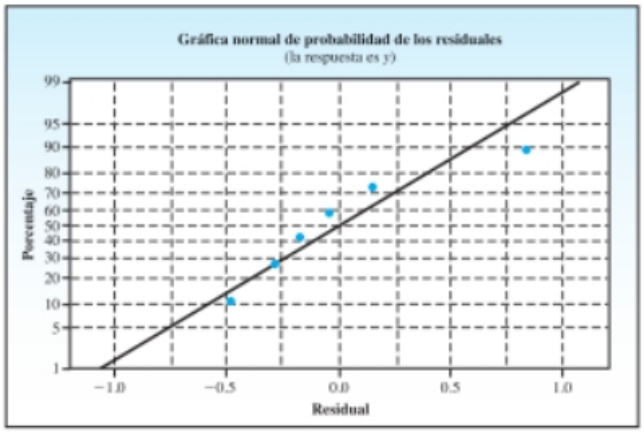
c. ¿Los datos dan suficiente evidencia para indicar que el modelo contribuye con información para la predicción de y? ¿Cuál es el valor p para esta prueba y qué significa?

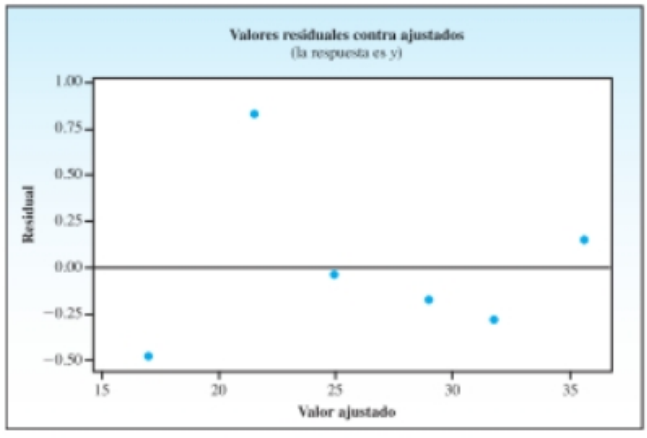
d. ¿Qué signo se esperaría que tenga el valor real de β_2? Encuentre el valor de β_2 en la salida impresa. ¿Este valor confirma las expectativas?

e. Los datos indican una curvatura significativa en la relación entre y y x? Pruebe al nivel de significancia de 5%.

f. ¿Qué conclusiones se pueden sacar de las gráficas residuales siguientes?

Gráficas de diagnóstico para el ejercicio 9.8





DATOS EX1315

9.9 Utilidades corporativas Para estudiar la relación de publicidad e inversión de capital con utilidades corporativas, los datos siguientes, registrados en unidades de 100 000 dólares, se recolectaron para 10 empresas de mediano tamaño en el mismo año. La variable y representa utilidad para el año, x_1 representa inversión de capital y x_2 representa gasto en publicidad.

y	x_1	x_2	y	x_1	x_2
15	25	4	1	20	0
16	1	5	16	12	4
2	6	3	18	15	5
3	30	1	13	6	4
12	29	2	2	16	2

a. Usando el modelo

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

y un paquete apropiado de software, encuentre la ecuación de predicción de mínimos cuadrados para estos datos.

b. Use la prueba F general para determinar si el modelo aporta información significativa para la predicción de y. Use $\alpha = .01$.

c. ¿El gasto en publicidad x_2 aporta información significativa para la predicción de y, dado que x_1 ya está en el modelo? Use $\alpha = .01$.

d. Calcule el coeficiente de determinación, R^2. ¿Qué porcentaje de la variación general está explicado por el modelo?

DATOS EX1326

9.10 Publicidad y ventas Una tienda de departamentos realizó un experimento con el fin de investigar los efectos de gastos de publicidad en las ventas semanales de sus departamentos de ropa para caballeros, para niños y para damas. Cinco semanas de observación se seleccionaron al azar de cada departamento y un presupuesto de publicidad x_1 (en cientos de dólares) se asignó a cada uno. Las ventas semanales (en miles de dólares) se muestran en la tabla siguiente para cada uno de los 15 periodos de venta de una semana. Si esperamos que las ventas semanales $E(y)$ estén linealmente relacionadas con el gasto de publicidad x_1 y si esperamos que las pendientes de las rectas correspondientes a los tres departamentos difieran, entonces un modelo apropiado para $E(y)$ es

$$E(y) = \underbrace{\beta_0 + \beta_1 x_1}_{\text{variable cuantitativa "gasto en publicidad"}} + \underbrace{\beta_2 x_2 + \beta_3 x_3}_{\text{variables ficticias usadas para introducir la variable cualitativa "departamento" en el modelo}} + \underbrace{\beta_4 x_1 x_2 + \beta_5 x_1 x_3}_{\text{términos de interacción que introducen diferencias en pendientes}}$$

donde

x_1 = gasto en publicidad

$$x_2 = \begin{cases} 1 & \text{si el departamento B de ropa para niños} \\ 0 & \text{si no lo es} \end{cases}$$

$$x_3 = \begin{cases} 1 & \text{si el departamento C de ropa para damas} \\ 0 & \text{si no lo es} \end{cases}$$

	Gasto en publicidad (cientos de dólares)				
Departamento	1	2	3	4	5
Ropa para caballeros A	5.2	5.9	7.7	7.9	9.4
Ropa para niños B	8.2	9.0	9.1	10.5	10.5
Ropa para damas C	10.0	10.3	12.1	12.7	13.6

a. Encuentre la ecuación de la recta que relacione $E(y)$ con el gasto en publicidad x_1 para el departamento A de ropa para caballeros. [SUGERENCIA: De acuerdo con el código usado para las variables ficticias, el modelo representa ventas medias $E(y)$ para el departamento A de ropa para caballeros cuando $x_2 = x_3 = 0$. Sustituya $x_2 = x_3 = 0$ en la ecuación para $E(y)$ para hallar la ecuación de esta recta.]

b. Encuentre la ecuación de la recta que relacione $E(y)$ con x_1 para el departamento B de ropa para niños. [SUGERENCIA: De acuerdo con el código, el modelo representa $E(y)$ para el departamento de ropa para niños cuando $x_2 = 1$ y $x_3 = 0$.]

c. Encuentre la ecuación para la recta que relacione $E(y)$ con x_1 para el departamento C de ropa para damas.

d. Encuentre la diferencia entre los puntos de intersección de las rectas $E(y)$ correspondientes a los departamentos B de ropa para niños y A de ropa para caballeros.

e. Encuentre la diferencia de pendientes entre las rectas $E(y)$ correspondientes a los departamentos C de ropa para damas y A de ropa para caballeros.

f. Consulte la parte **e**. Suponga que se desea probar la hipótesis nula de que las pendientes de las rectas correspondientes a los tres departamentos son iguales. Exprese esto como una prueba de hipótesis acerca de uno o más parámetros del modelo.

10

Series de tiempo

© Bob Pardue-Medical Lifestyle/Alamy.

OBJETIVOS GENERALES

El propósito de este capítulo es presentar el análisis de series de tiempo y la elaboración de pronósticos. Los métodos de elaboración de pronósticos se pueden clasificar como cualitativos o cuantitativos. Debido a que este capítulo se enfoca en los métodos de series de tiempo, se deja la discusión acerca de la aplicación del análisis de regresión como método de elaboración de pronósticos causal a textos más avanzados sobre la materia.

ÍNDICE DEL CAPÍTULO

- Componentes de la serie de tiempo (tipo de variaciones)
- Análisis de tendencias
- Variación estacional
- Variación cíclica
- Variación irregular
- Problemas que incluyen los cuatro componentes
- Análisis de las series de tiempo en los pronósticos

10.1 COMPONENTES DE LA SERIE DE TIEMPO

Una **serie de tiempo** es una secuencia de observaciones en una variable que se mide en puntos sucesivos en el tiempo o sobre periodos sucesivos. Las medidas pueden tomarse cada hora, día, semana, mes o año, o en cualquier otro intervalo regular.[1] La variación de datos es un factor importante en la comprensión de cómo se han comportado las series de tiempo en el pasado. Si se espera que tal comportamiento continúe en el futuro, se puede utilizar la variación anterior como guía en la selección de un método adecuado de elaboración de pronósticos.

Para identificar los datos de la variación subyacente, un primer paso útil es elaborar una **gráfica de series de tiempo**. Ésta es una representación gráfica de la relación entre el tiempo y las variables de serie de tiempo: el tiempo se encuentra en el eje horizontal y los valores de la serie de tiempo en el eje vertical. Se revisarán algunos tipos comunes de variaciones de datos que se pueden identificar al analizar una gráfica de series de tiempo.

Variación horizontal

Una **variación horizontal** tiene lugar cuando los datos fluctúan alrededor de una media constante. Para ilustrar una serie de tiempo con una variación horizontal, observe los datos de la tabla 10.1.

TABLA 10.1 **Serie de tiempo de las ventas de gasolina**

Semana	Ventas (miles de galones)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22

WEB file
Gasoline

Estos datos muestran el número de galones de gasolina vendidos por un distribuidor en Bennington, Vermont, en las últimas 12 semanas. El valor medio o promedio de esta serie de tiempo es 19.25 o 19 250 galones por semana. La figura 10.1 muestra una gráfica de serie de tiempo para estos datos. Observe cómo éstos fluctúan alrededor de una media muestral de 19 250 galones. Aunque la variabilidad aleatoria está presente, se diría que estos datos siguen una variación horizontal.

El concepto de **series de tiempo estacionarias**[2] designa una serie de tiempo cuyas propiedades estadísticas son independientes del tiempo. Esto significa, en particular, que

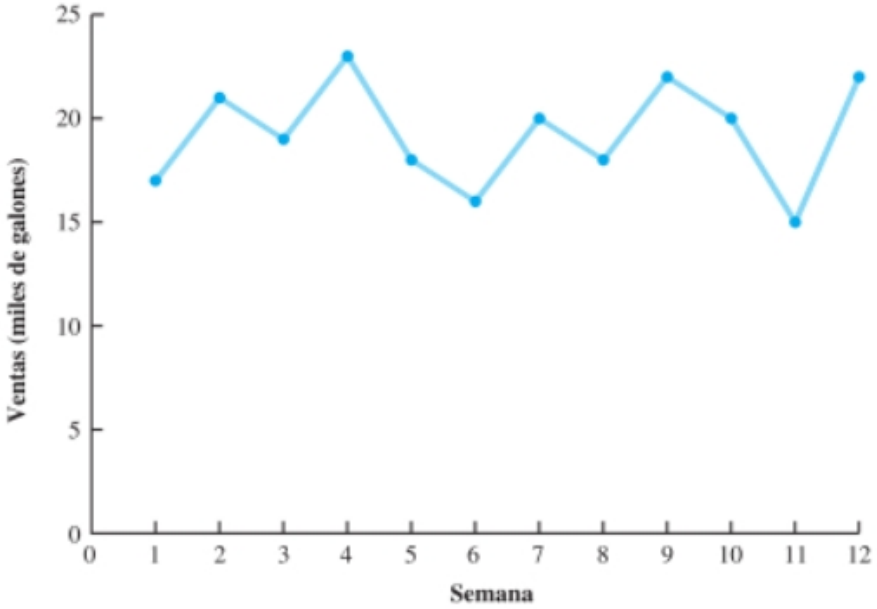
1. El proceso de generación de datos tiene una media constante.
2. La variabilidad de la serie de tiempo es constante en el tiempo.

Una gráfica de una serie de tiempo estacionaria exhibe siempre una variación horizontal. Pero la sola observación de una variación horizontal no es evidencia suficiente para concluir que la serie de tiempo sea estacionaria. Los libros más avanzados sobre elaboración de pronósticos

[1] El estudio se limitará a series de tiempo en las que los valores de las series se miden en intervalos iguales. Los casos en los que las observaciones se realizan en intervalos desiguales van más allá del alcance de este texto.

[2] Para una definición formal de series de tiempo estacionarias, remítase a G. E. P. Box, G. M. Jenkins y G. C. Reinsell, *Time Series Analysis: Forecasting and Control* (*Análisis de series de tiempo: pronóstico y control*), 3a. ed., Englewood Cliffs, NJ, Prentice Hall, 1994, p. 23.

FIGURA 10.1
Gráfica de la serie de tiempo de las ventas de gasolina



estudian los procedimientos para determinar si una serie de tiempo es estacionaria, y proporcionan métodos para transformarla de no estacionaria en estacionaria.

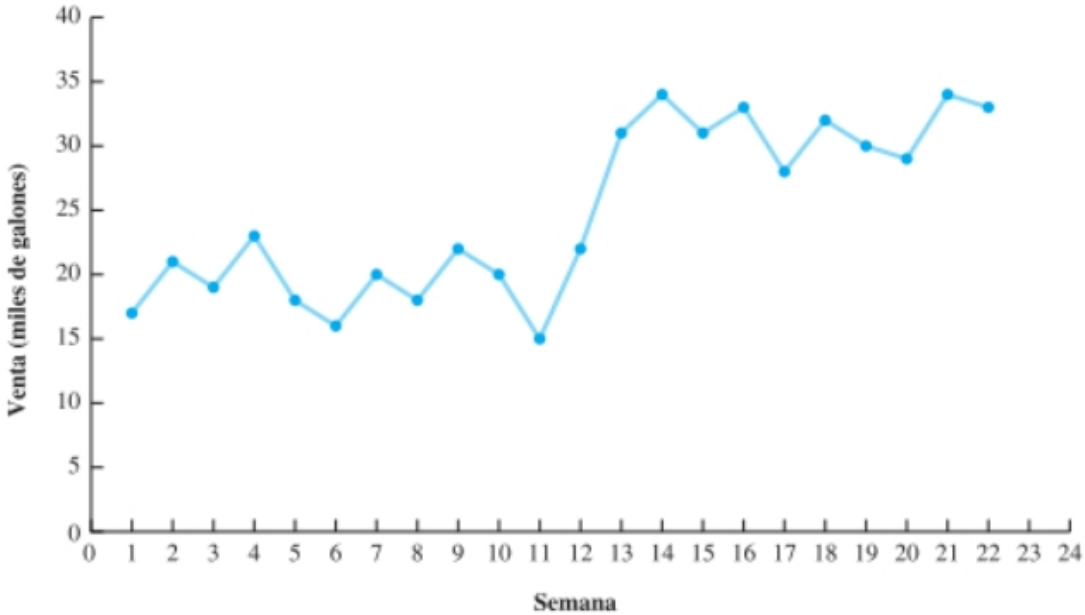
Los cambios en las condiciones de negocios a menudo pueden dar lugar a que una serie de tiempo que tiene una variación horizontal cambie a un nuevo nivel. Por ejemplo, suponga que el distribuidor de gasolina firma un contrato con el Departamento de Policía de Vermont para proveer de gasolina a las patrullas de la policía local ubicadas al sur del estado. Con este nuevo contrato el distribuidor espera tener un gran incremento en las ventas semanales a partir de la semana 13. La tabla 10.2 muestra el número de galones de gasolina que se venden para la serie de tiempo original y para las 10 semanas después de firmar el nuevo contrato. La figura 10.2 muestra la gráfica correspondiente de series de tiempo. Observe el incremento en el nivel de la serie de tiempo a partir de la semana 13. Este cambio hace más difícil elegir un método adecuado de elaboración de pronósticos. La selección de un método que se adapte bien a los cambios en el nivel de una serie de tiempo es una consideración importante en muchas aplicaciones prácticas.

TABLA 10.2
Serie de tiempo de las ventas de gasolina después de obtener el contrato con la policía de Vermont

Semana	Ventas (miles de galones)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22
13	31
14	34
15	31
16	33
17	28
18	32
19	30
20	29
21	34
22	33

FIGURA 10.2

Gráfica de la serie de tiempo de ventas de gasolina después de obtener el contrato con la policía de Vermont



Variación de tendencia

Aunque los datos de las series de tiempo presentan fluctuaciones aleatorias, estas series también pueden mostrar cambios o movimientos graduales hacia valores relativamente mayores o menores durante un periodo. Si una gráfica de series de tiempo muestra este tipo de comportamiento, se dice que existe una **variación de tendencia**. La tendencia, por lo general, es el resultado de factores a largo plazo, como el aumento o disminución de la población o la variación de sus características demográficas, la tecnología y/o preferencias de los consumidores, etcétera.

Para ilustrar una serie de tiempo con una variación de tendencia, considere las series de tiempo de ventas de bicicletas de cierto fabricante en los últimos 10 años, como se muestra en la tabla 10.3 y la figura 10.3. Observe que en el primer año se vendieron 21 600 bicicletas, en el segundo 22 900, y así sucesivamente. En el año 10, el último, se han vendido 31 400 bicicletas. La inspección visual de la gráfica de la serie de tiempo permite apreciar algunos movimientos ascendentes y descendentes en los últimos 10 años, pero la serie de tiempo también parece mostrar una tendencia sistemática de aumento o disminución.

La tendencia en la serie de tiempo de ventas de bicicletas parece ser lineal y aumentar con el tiempo, pero a veces una tendencia se puede describir mejor por otros tipos de variaciones. Por ejemplo, los datos en la tabla 10.4 y la gráfica correspondiente a la serie de tiempo de la figura 10.4 muestran las ventas de un medicamento contra el colesterol, dado que la empresa obtuvo la aprobación de la FDA hace 10 años. La serie de tiempo se incrementó de forma no lineal, es decir, la tasa de variación de los ingresos no aumentó una cantidad constante de un año a otro. De hecho, los ingresos parecen estar creciendo de manera exponencial. Las relaciones exponenciales de este tipo son apropiadas cuando la variación porcentual de un periodo a otro es relativamente constante.

TABLA 10.3 **Serie de tiempo de ventas de bicicletas**

Año	Ventas (miles)
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4

TABLA 10.4

Serie de tiempo de ingresos por medicamentos contra el colesterol (millones de $)

Año	Ingresos
1	23.1
2	21.3
3	27.4
4	34.6
5	33.8
6	43.2
7	59.5
8	64.4
9	74.2
10	99.3

Variación o componente estacional

La tendencia de una serie de tiempo se puede identificar con el análisis de las variaciones multianuales en los datos históricos. Las variaciones estacionales son reconocidas al identificarse las mismas variaciones de repetición en periodos sucesivos. Por ejemplo, un fabricante de albercas espera tener pocas ventas en los meses de otoño e invierno, y aumentarlas en los meses de primavera y verano. Los fabricantes de equipos de remoción de nieve y de ropa de invierno, sin embargo, prevén exactamente lo contrario. Como era de esperar, la variación de una gráfica de series de tiempo que tiene un comportamiento repetitivo en un periodo de un año debido a la influencia estacional se llama **variación o componente estacional**. Aunque por lo general se considera que las variaciones estacionales son aquellas que se presentan en un lapso de un año, los datos de series de tiempo también pueden presentar variaciones estacionales de menos de un año. Por ejemplo, el volumen de tráfico diario muestra un día un comportamiento "estacional", donde los valores máximos se presentan en las horas de máxima afluencia de vehículos, un flujo moderado el resto del día y al comienzo de la noche, y un flujo ligero desde la medianoche hasta la madrugada.

Como ejemplo de variación estacional, considere el número de sombrillas vendidas en una tienda de ropa en los últimos cinco años. La tabla 10.3 muestra la serie de tiempo con los datos de año (Year), trimestre (Quarter) y ventas (Sales), y la figura 10.5 ilustra la gráfica correspondiente. La gráfica de la serie de tiempo no indica tendencia alguna a largo plazo en las ventas. De hecho, a menos que observe cuidadosamente los datos, es posible concluir que éstos siguen una variación horizontal. Pero una inspección más cercana revela una variación regular en los datos. Es decir, el primer y el tercer trimestres presentan ventas moderadas, el

FIGURA 10.3
Gráfica de la serie de tiempo de las ventas de bicicletas

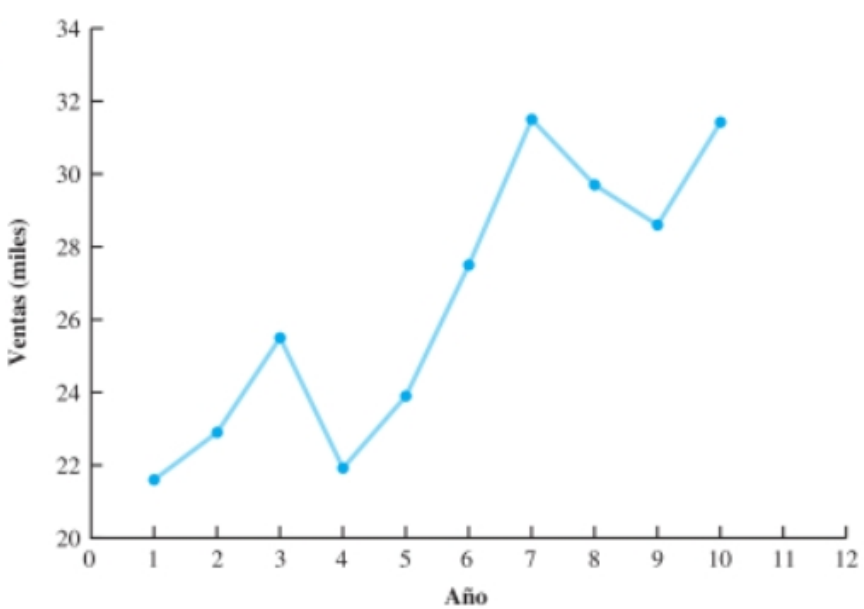
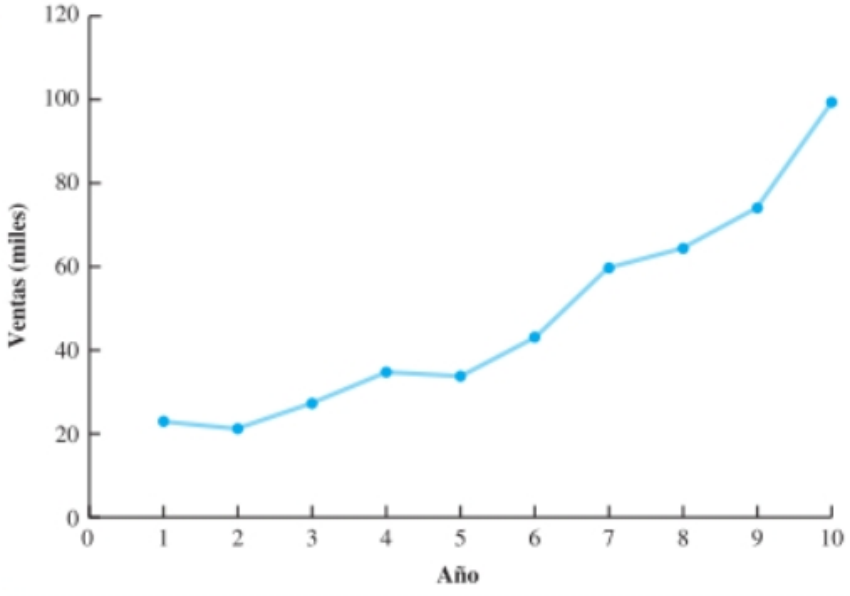


FIGURA 10.4
Gráfica de la serie de tiempo de ventas de medicamentos contra el colesterol (millones de $)



segundo trimestre tiene ventas más altas, y el cuarto trimestre tiende a tener el menor volumen de ventas. Por lo tanto, se concluye que existe una variación estacional trimestral.

Variaciones estacional y de tendencia

Algunas series de tiempo son una combinación de una tendencia y una variación estacional. Por ejemplo, los datos de la tabla 10.6 y la gráfica correspondiente de la serie de tiempo en la figura 10.6 muestran las ventas (Sales) de televisores por trimestre (Quarter) y año (Year) de un fabricante en los últimos cuatro años. Claramente se presenta una tendencia creciente. Sin embargo, la figura 10.6 indica también que las ventas son menores en el segundo trimestre de cada año y que aumentan a partir de los trimestres 3 y 4. Por lo tanto, se llega a la conclusión de que una variación estacional también se encuentra en las ventas de televisores. En estos casos se utiliza un método de elaboración de pronósticos que tiene la capacidad de tratar la tendencia y la estacionalidad.

TABLA 10.5 **Serie de tiempo de ventas de sombrillas**

WEB file
Umbrella

Year	Quarter	Sales
1	1	125
	2	153
	3	106
	4	88
2	1	118
	2	161
	3	133
	4	102
3	1	138
	2	144
	3	113
	4	80
4	1	109
	2	137
	3	125
	4	109
5	1	130
	2	165
	3	128
	4	96

FIGURA 10.5

Gráfica de la serie de tiempo de ventas de sombrillas

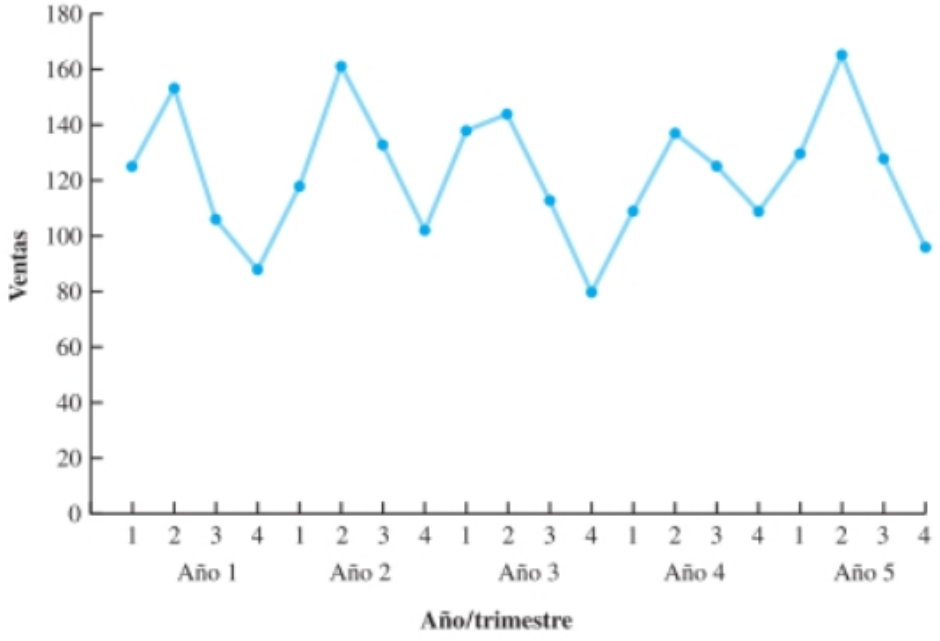


TABLA 10.6

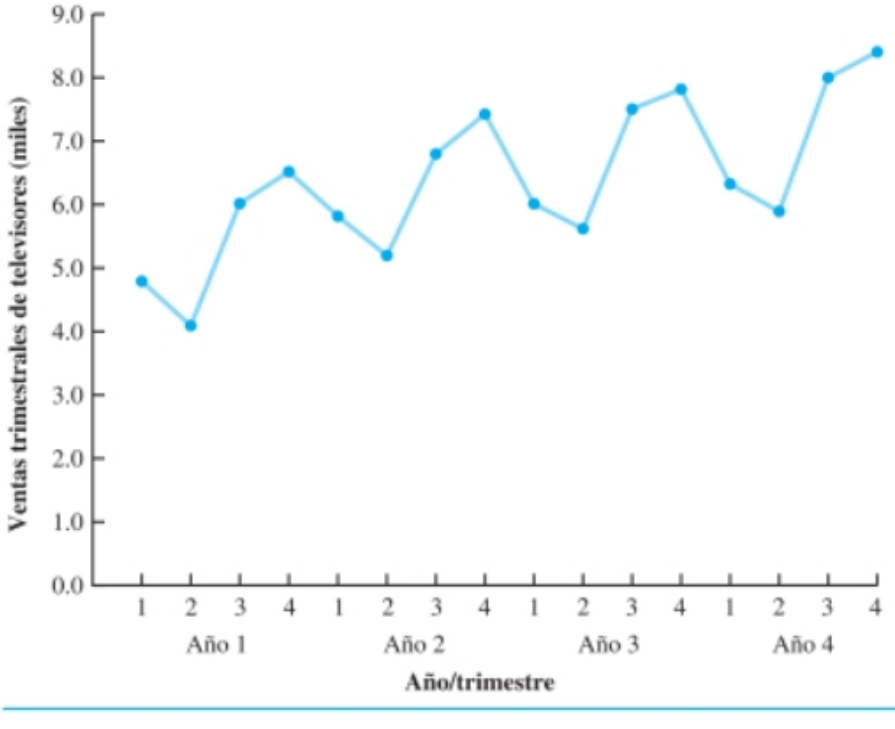
Serie de tiempo de las ventas de televisores

Year	Quarter	Sales (1000s)
1	1	4.8
	2	4.1
	3	6.0
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6.0
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8.0
	4	8.4

WEB file
TVSales

FIGURA 10.6

Gráfica de la serie de tiempo de ventas trimestrales de televisores



10.2 ANÁLISIS DE TENDENCIAS

WEB file
Bicycle

En esta sección se presentan tres métodos de elaboración de pronósticos que son apropiados para las series de tiempo que exhiben una variación de tendencia. En primer lugar, se muestra qué tan sencillo es aplicar la regresión lineal para pronosticar la serie de tiempo con tendencia lineal. Después se ilustra cómo obtener los pronósticos mediante el suavizamiento exponencial lineal de Holt, una extensión del único suavizamiento exponencial que utiliza dos constantes de suavizamiento: la primera se toma en cuenta para el nivel de la serie de tiempo y la segunda para la tendencia lineal de los datos. Por último, se muestra cómo la capacidad de ajuste a una curva en el análisis de regresión permite pronosticar series de tiempo con una tendencia curvilínea o no lineal.

Regresión de tendencia lineal

En la sección 10.1 se utilizaron las ventas de bicicletas que se incluyen en la tabla 10.3 y en la figura 10.3 para ilustrar una serie de tiempo con una variación de tendencia. Esta serie de tiempo se utilizará para ilustrar cómo la regresión lineal simple permite pronosticar una serie de tiempo con tendencia lineal. Los datos de las series de tiempo de las ventas de bicicletas se repiten en la tabla 10.12 y en la figura 10.7.

Aunque la gráfica de la serie de tiempo de la figura 10.7 muestra algún movimiento ascendente y descendente en los últimos 10 años, se podría estar de acuerdo en que la línea de tendencia que se muestra en la figura 10.8 proporciona una aproximación razonable del movimiento de la serie a largo plazo. Se pueden utilizar los métodos de regresión lineal simple a efecto de obtener una tendencia lineal para las series de tiempo de ventas de bicicletas.

TABLA 10.7 **Serie de tiempo de ventas de bicicletas**

Año	Ventas (en miles)
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4

FIGURA 10.7
Gráfica de la serie de tiempo de ventas de bicicletas

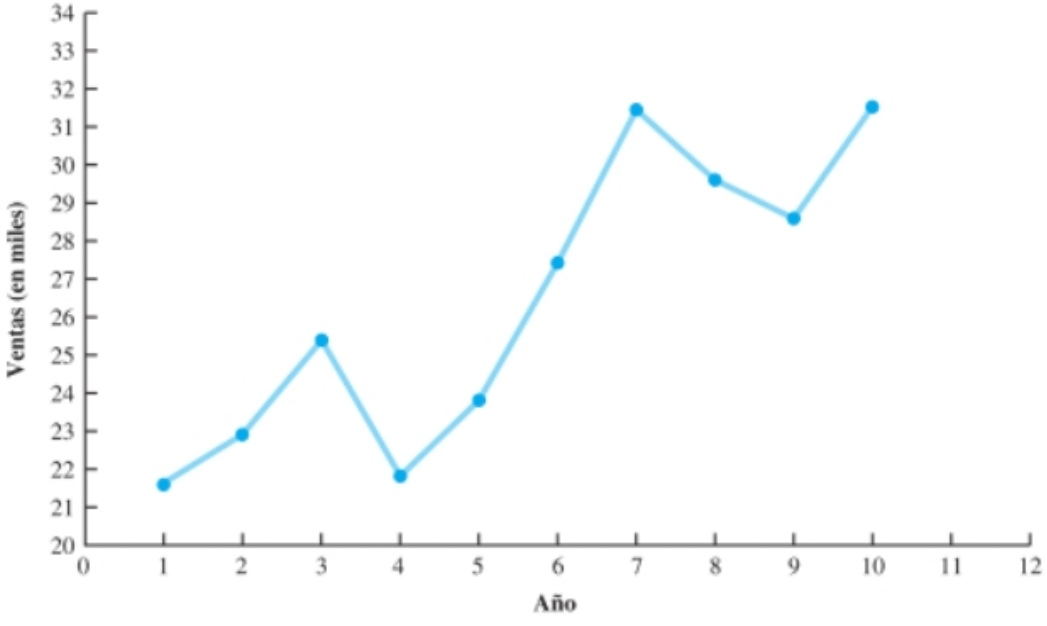
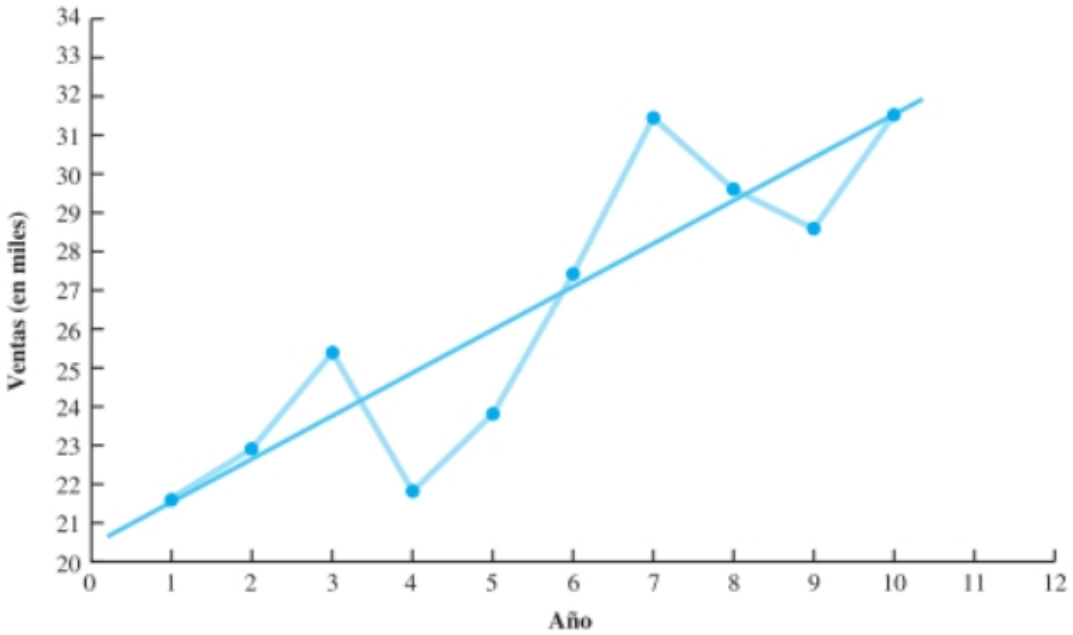


FIGURA 10.8
Tendencia representada por una función lineal de la serie de tiempo de ventas de bicicletas



La ecuación de regresión estimada que describe una relación lineal entre una variable independiente x y una variable dependiente y se expresa como

$$\hat{y} = b_0 + b_1 x$$

donde $\hat{y}$ es el valor estimado o predicho de y. Para enfatizar el hecho de que en un pronóstico la variable independiente es el tiempo, se sustituirá x por t, así como $\hat{y}$ por T_t para hacer patente que se está estimando la tendencia de una serie de tiempo. Por lo tanto, para calcular la tendencia lineal en una serie de tiempo se utilizará la siguiente ecuación de regresión estimada.

ECUACIÓN DE TENDENCIA LINEAL

$$T_t = b_0 + b_1 t \tag{10.1}$$

donde

T_t = pronóstico de tendencia lineal en el periodo t

b_0 = intersección de la línea de tendencia lineal

b_1 = pendiente de la línea de tendencia lineal

t = periodo

En la ecuación (10.1) la variable tiempo comienza en $t = 1$ correspondiente a la primera observación de la serie de tiempo (año 1 para la serie de tiempo de ventas de bicicletas) y continúa hasta que $t = n$ correspondiente a la observación más reciente (año 10 para la serie de tiempo de ventas de bicicletas). Así, $t = 1$ corresponde al valor de la serie más antigua y $t = 10$ corresponde al año más reciente.

A continuación se presentan las fórmulas para calcular los coeficientes de regresión estimados b_0 y b_1 de la ecuación (10.1).

CÁLCULO DE LA PENDIENTE Y DE LA INTERSECCIÓN PARA UNA TENDENCIA LINEAL*

$$b_1 = \frac{\sum_{t=1}^{n}(t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(t - \bar{t})^2} \tag{10.2}$$

$$b_0 = \bar{Y} - b_1\bar{t} \tag{10.3}$$

donde

Y_t = valor de la serie de tiempo en el periodo t

n = número de periodos (número de observaciones)

$\bar{Y}$ = valor promedio de la serie de tiempo

$\bar{t}$ = valor promedio de t

*Otra fórmula para b_1 es

$$b_1 = \frac{\sum_{t=1}^{n} tY_t - \left(\sum_{t=1}^{n} t \sum_{t=1}^{n} Y_t\right)/n}{\sum_{t=1}^{n} t^2 - \left(\sum_{t=1}^{n} t\right)^2/n}$$

Esta forma de la ecuación (10.2) se recomienda a menudo cuando se utiliza una calculadora para obtener b_1.

A efecto de calcular la ecuación de tendencia lineal para la serie de tiempo de las ventas de bicicletas, se comienza por determinar $\bar{t}$ y $\bar{Y}$ utilizando la información de la tabla 10.7.

$$\bar{t} = \frac{\sum_{t=1}^{n} t}{n} = \frac{55}{10} = 5.5$$

$$\bar{Y} = \frac{\sum_{t=1}^{n} Y_t}{n} = \frac{264.5}{10} = 26.45$$

TABLA 10.8 **Resumen de los cálculos de la tendencia lineal para la serie de tiempo de ventas de bicicletas**

	3	25.5	−2.5	−0.95	2.375	6.25
	4	21.9	−1.5	−4.55	6.825	2.25
	5	23.9	−0.5	−2.55	1.275	0.25
	6	27.5	0.5	1.05	0.525	0.25
	7	31.5	1.5	5.05	7.575	2.25
	8	29.7	2.5	3.25	8.125	6.25
	9	28.6	3.5	2.15	7.525	12.25
	10	31.4	4.5	4.95	22.275	20.25
Totales	55	264.5			90.750	82.50

Con estos valores y la información de la tabla 10.8 se puede calcular la pendiente y la intersección de la tendencia lineal para la serie de tiempo de ventas de bicicletas.

$$b_1 = \frac{\sum_{t=1}^{n}(t - \bar{t})(Y_t - \bar{Y})}{\sum_{t=1}^{n}(t - \bar{t})^2} = \frac{90.75}{82.5} = 1.1$$

$$b_0 = \bar{Y} - b_1\bar{t} = 26.45 - 1.1(5.5) = 20.4$$

Por lo tanto, la ecuación de tendencia lineal es

$$T_t = 20.4 = 1.1t$$

La pendiente de 1.1 indica que en los últimos 10 años la empresa experimentó un promedio de crecimiento en las ventas de cerca de 1 100 unidades por año. Si se supone que la tendencia en las ventas de la última década es un buen indicador del futuro, esta ecuación de tendencia puede utilizarse para formular los pronósticos sobre periodos posteriores. Por ejemplo, al sustituir en la ecuación $t = 11$ se obtiene la proyección de tendencia o el pronóstico para el próximo año T_{11}.

$$T_{11} = 20.4 \lesseqgtr 1.1(11) = 32.59$$

Así, al utilizar la proyección de tendencia se podrá pronosticar un valor para las ventas de 32 500 bicicletas para el próximo año.

Para calcular la exactitud del método de elaboración de pronósticos de la proyección de tendencia se utilizará el CME. La tabla 10.9 muestra el cálculo de la suma de los errores cuadrados para la serie de tiempo de ventas de bicicletas. Así que para esta serie,

$$\text{CME} = \frac{\sum_{t=1}^{n}(Y_t - F_t)^2}{n} = \frac{30.7}{10} = 3.07$$

Debido a que el pronóstico de regresión de tendencia lineal utiliza el mismo procedimiento del análisis de regresión, se pueden utilizar los procedimientos estándar para realizar los cálculos del análisis de regresión con *MINITAB* o Excel. La figura 10.9 muestra el resultado obtenido al utilizar el módulo de *MINITAB* para el análisis de regresión de la serie de tiempo de ventas de bicicletas.

En la figura 10.9 el valor del CME en la tabla de ANOVA es

$$\text{CME} = \frac{\text{Suma de cuadrados debido al error}}{\text{Grados de libertad}} = \frac{30.7}{8} = 3.837$$

TABLA 10.9 **Resumen de los pronósticos de tendencia lineal y de errores de pronóstico de la serie de tiempo de ventas de bicicletas**

Año	Ventas (en miles) Y_t	Pronóstico T_t	Error de pronóstico	Error cuadrado de pronóstico
1	21.6	21.5	0.1	0.01
2	22.9	22.6	0.3	0.09
3	25.5	23.7	1.8	3.24
4	21.9	24.8	−2.9	8.41
5	23.9	25.9	−2.0	4.00
6	27.5	27.0	0.5	0.25
7	31.5	28.1	3.4	11.56
8	29.7	29.2	0.5	0.25
9	28.6	30.3	−1.7	2.89
10	31.4	31.4	0.0	0.00
			Total	30.70

En los resultados de MINITAB del análisis de tendencia, MSD *es la desviación cuadrada media, es decir, el promedio de los errores cuadrados de pronóstico.*

Este valor del CME difiere del valor calculado antes, porque la suma de los errores cuadrados se divide entre 8 en vez de 10; por lo tanto, en el CME el resultado de regresión no es la media de los errores cuadrados de pronóstico. Sin embargo, la mayoría del software de pronósticos calculan el valor del CME tomando la media de los errores cuadrados. Por lo tanto, al utilizar los paquetes de series de tiempo para desarrollar una ecuación de tendencia, el resultado del valor del CME puede diferir ligeramente del que se obtendría con un método general de regresión. Por ejemplo, en la figura 10.10 se muestra la parte gráfica del resultado obtenido al utilizar el procedimiento de análisis de tendencia de series de tiempo de *MINITAB*. Observe que MSD = 3.07 es el promedio de los errores cuadrados de pronóstico.

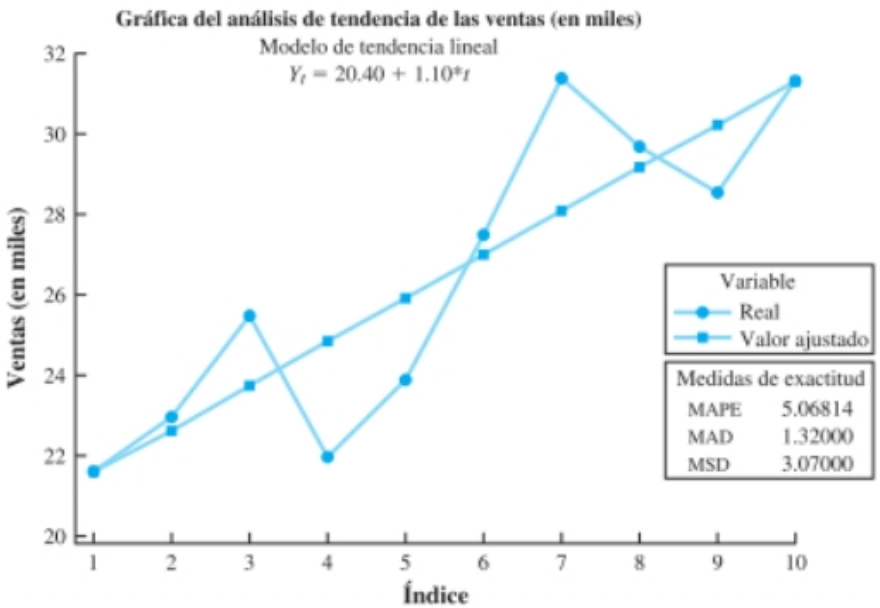
FIGURA 10.9

Resultado de la regresión de *MINITAB* de la serie de tiempo de ventas de bicicletas

```
The regression equation is
Y = 20.4 + 1.10 t

Predictor     Coef   SE Coef       T       p
Constant    20.400     1.338   15.24   0.000
t           1.1000    0.2157    5.10   0.001

S = 1.95895   R-sq = 76.5%   R-sq(adj) = 73.5%

Analysis of Variance

SOURCE           DF       SS      MS      F      p
Regression        1   99.825  99.825  26.01  0.001
Residual Error    8   30.700   3.837
Total             9  130.525
```

FIGURA 10.10

Resultado del análisis de tendencia lineal en *MINITAB* de la serie de tiempo de ventas de bicicletas



Suavizamiento exponencial lineal de Holt

El suavizamiento exponencial lineal de Holt suele llamarse **suavizamiento exponencial doble**.

Charles Holt desarrolló una versión de suavizamiento exponencial para pronosticar una serie de tiempo con tendencia lineal. Recuerde que en el procedimiento de suavizamiento exponencial se utiliza α como constante de suavizamiento para "suavizar" la aleatoriedad o las fluctuaciones irregulares en una serie de tiempo; los pronósticos para el periodo $t = 1$ se obtienen mediante la ecuación

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

Los pronósticos por el método de **suavizamiento exponencial lineal de Holt** se obtienen mediante dos constantes de suavizamiento, α y β, y tres ecuaciones.

ECUACIONES DE SUAVIZAMIENTO EXPONENCIAL LINEAL DE HOLT

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (10.4)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)\, b_{t-1} \quad (10.5)$$

$$F_{t+k} = L_t + b_t\, k \quad (10.6)$$

donde

L_t = estimación del nivel de la serie de tiempo para el periodo t
b_t = estimación de la pendiente de la serie de tiempo para el periodo t
α = constante de suavizamiento para el nivel de la serie de tiempo
β = constante de suavizamiento para la pendiente de la serie de tiempo
F_{t+k} = pronóstico para el periodo k en el futuro
k = número de periodos próximos a pronosticar

Aplique el método de Holt a la serie de tiempo de las ventas de bicicletas listadas en la tabla 10.7 y utilice $\alpha = 0.1$ y $\beta = 0.2$. Para iniciar con el método se necesitan los valores de L_1, la estimación del nivel de la serie de tiempo en el año 1, y b_1, la estimación de la pendiente de la serie de tiempo en el año 1. El método que se utiliza por lo general consiste en determinar $L_1 = Y_1$ y $b_1 = Y_2 - Y_1$. Al utilizar este procedimiento inicial obtenemos

$$L_1 = Y_1 = 21.6$$
$$b_1 = Y_2 - Y_1 = 22.9 - 21.6 = 1.3$$

Mediante la ecuación (10.6) con $k = 1$, el pronóstico de ventas en el año 2 es $F_2 = L_1 + b_1 = 21.6 \lesssim 1.3(1) = 22.9$. Después se continúa utilizando las ecuaciones (10.4) a (10.6) para calcular las estimaciones del nivel y la tendencia para el año 2, así como un pronóstico para el año 3.

Primero utilizamos la ecuación (10.4) y la constante de suavizamiento $\alpha = 0.1$ para calcular la estimación del nivel de la serie de tiempo para el año 2.

$$L_2 = 0.1(22.9) \lesssim 0.9(21.6 \lesssim 1.3) = 22.9$$

Observe que $21.6 \lesssim 1.3$ es el pronóstico de ventas para el año 2. Por lo tanto, la estimación del nivel de la serie de tiempo en este año que se obtiene mediante la ecuación (10.4) es simplemente un promedio ponderado del valor observado en el año 2 (con un peso de $\alpha = 0.1$) y el pronóstico para el año 2 (con un peso de $1 - \alpha = 1 - 0.1 = 0.9$). En general, los valores mayores que α dan más peso al valor observado (Y_t), mientras que valores menores dan más peso al valor pronosticado ($L_{t-1} - b_{t-1}$).

A continuación utilice la ecuación (10.5) y la constante de suavizamiento $\beta = 0.2$ para calcular la pendiente de la serie de tiempo en el año 2.

$$b_2 = 0.2(22.9 - 21.6) - (1 - 0.2)(1.3) = 1.3$$

La estimación de tal pendiente es un promedio ponderado de la diferencia en el nivel estimado de la serie de tiempo entre el año 2 y el año 1 (con un peso de $\beta = 0.2$) y la estimación de la pendiente del año 1 (con un peso de $1 - \beta = 1 - 0.2 = 0.8$). En general, los valores mayores que β dan más peso a la diferencia de los niveles estimados, mientras que los valores menores se lo conceden a la pendiente estimada del último periodo.

Al utilizar las estimaciones de L_2 y b_2 recién obtenidas, el pronóstico de ventas para el año 3 se calcula mediante la ecuación (10.6).

$$F_3 = L_2 - b_2 = 22.9 - 1.3(1) = 24.2$$

De forma similar se realizan otros cálculos, los cuales se muestran en la tabla 10.10. La suma de los errores cuadrados de pronóstico es 39.678; por ende, CME = 39.678/9 = 4.41.

TABLA 10.10 **Resumen de los cálculos de Holt para el suavizamiento exponencial lineal de la serie de tiempo de ventas de bicicletas utilizando $\alpha = 0.1$ y $\beta = 0.2$**

Año	Ventas (miles) Y_t	Nivel estimado L_t	Tendencia estimada b_t	Pronóstico F_t	Error de pronóstico	Error cuadrado de pronóstico
1	21.6	21.600	1.300			
2	22.9	22.900	1.300	22.900	0.000	0.000
3	25.5	24.330	1.326	24.200	1.300	1.690
4	21.9	25.280	1.251	25.656	−3.756	14.108
5	23.9	26.268	1.198	26.531	−2.631	6.924
6	27.5	27.470	1.199	27.466	0.034	0.001
7	31.5	28.952	1.256	28.669	2.831	8.016
8	29.7	30.157	1.245	30.207	−0.507	0.257
9	28.6	31.122	1.189	31.402	−2.802	7.851
10	31.4	32.220	1.171	32.311	−0.911	0.830
					Total	39.678

¿Con valores diferentes para las constantes de suavizamiento α y β se obtienen pronósticos más precisos? Para responder esta pregunta habría que probar diferentes combinaciones de α y β para determinar si se puede encontrar una combinación con la que se obtenga un valor del CME menor que 4.41, el valor obtenido utilizando las constantes de suavizamiento $\alpha = 0.1$ y $\beta = 0.2$. Se puede realizar la búsqueda por ensayo y error de buenos valores α y β o mediante software de estadística más avanzado que tenga la opción de seleccionar un conjunto óptimo de constantes de suavizamiento.

Observe que la estimación del nivel de la serie de tiempo en el año 10 es $L_1 = 32.220$ y de la pendiente en el mismo año es $b_1 = 1.171$. Si se asume que la tendencia de los últimos 10 años en las ventas es un buen indicador del futuro, la ecuación (10.6) puede utilizarse para desarrollar pronósticos para periodos futuros. Por ejemplo, al sustituir $t = 11$ en la ecuación (10.6) se obtiene la proyección de tendencia para el próximo año o el pronóstico, F_{11}.

$$F_{11} = L_{10} = b_{10}(1) = 32.220 + 1.171 = 33.391$$

Por lo tanto, al utilizar el suavizamiento exponencial lineal de Holt se pronostica que el año próximo las ventas serán de 33 391 bicicletas.

Regresión de tendencia no lineal

Para modelar una tendencia suele utilizarse una función lineal. Sin embargo, como ya se vio, algunas veces las series de tiempo tienen tendencias curvilíneas o no lineales. Como ejemplo, considere los ingresos anuales en millones de dólares para las ventas de un medicamento contra el colesterol en los primeros 10 años. La tabla 10.11 muestra la serie de tiempo y la figura 10.10 la gráfica correspondiente. Por ejemplo, los ingresos del año 1 fueron de \$23.1 millones; los del año 2 es \$21.3 millones, y así sucesivamente. La gráfica de la serie de tiempo indica una tendencia general creciente o a la alza. Pero a diferencia de la serie de tiempo de las ventas de bicicletas, la tendencia lineal no parece adecuada. En cambio, parece requerirse una función curvilínea para modelar la tendencia a largo plazo.

Ecuación de tendencia cuadrática

Diversas funciones no lineales pueden utilizarse para obtener una estimación de la tendencia de la serie de tiempo del colesterol. Por ejemplo, considere la siguiente ecuación de tendencia cuadrática:

WEB file
Cholesterol

$$T_t = b_0 + b_1 t = b_2 t^2 \qquad (10.7)$$

Para la serie de tiempo del colesterol, $t = 1$ corresponde al año 1, $t = 2$ al año 2, y así sucesivamente.

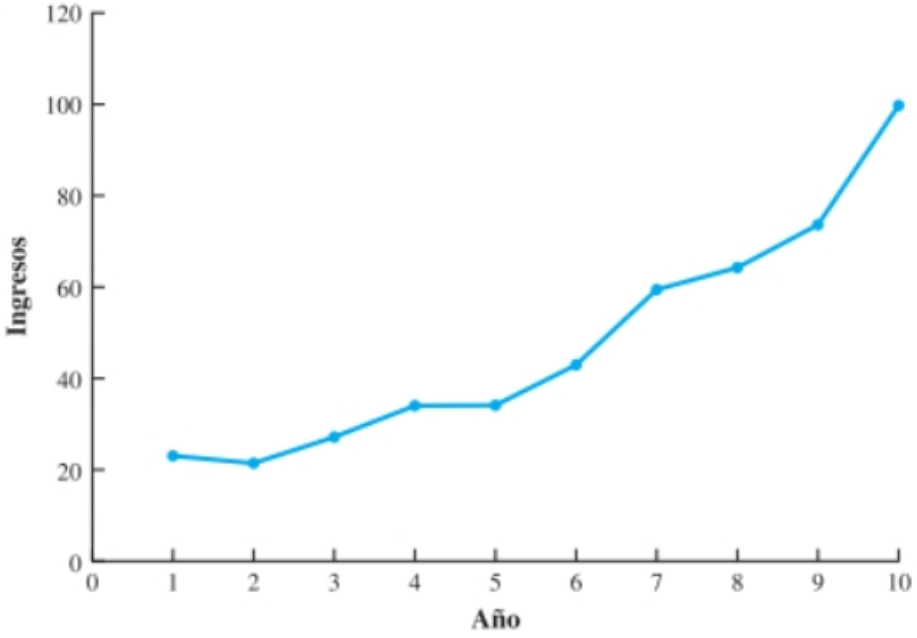
El modelo lineal general puede utilizarse para calcular los valores de b_0, b_1 y b_2. Existen dos variables independientes, el año y el año cuadrado, mientras que la variable dependiente son las ventas en millones de dólares. Así que la primera observación es 1, 1, 23.1; la segunda es 2, 4, 21.3; la tercera es 3, 9, 27.4, y así sucesivamente. La figura 10.12 muestra el resultado de *MINITAB* de la regresión múltiple para el modelo de tendencia cuadrática; la ecuación de regresión estimada es

$$\text{Revenue} = 24.2 - 2.11\ \text{Year} + 0.922\ \text{YearSq}.$$

TABLA 10.11 **Serie de tiempo de ventas de medicamentos contra el colesterol (millones de \$)**

Año (t)	Ingresos (millones de \$)
1	23.1
2	21.3
3	27.4
4	34.6
5	33.8
6	43.2
7	59.5
8	64.4
9	74.2
10	99.3

FIGURA 10.11
Gráfica de la serie de tiempo de ventas del medicamento contra el colesterol (millones de $)



donde

$$\text{Year} = 1, 2, 3, \ldots, 10$$

$$\text{YearSq} = 1, 4, 9, \ldots, 100$$

Al utilizar el procedimiento estándar de regresión múltiple se requiere calcular los valores para el año cuadrado como segunda variable independiente. Se puede utilizar como alternativa el análisis de tendencia para series de tiempo de *MINITAB* para obtener los mismos resultados. No se requiere obtener valores para el año cuadrado y es más fácil de usar. Se recomienda este método en la solución de los ejercicios que involucren tendencias cuadráticas.

Ecuación de tendencia exponencial

Otra alternativa que podemos utilizar para modelar la variación no lineal mostrada por la serie de tiempo del colesterol es ajustar a los datos un modelo exponencial. Por ejemplo, considere la siguiente ecuación de tendencia exponencial:

$$T_t = b_0(b_1)^t \tag{10.8}$$

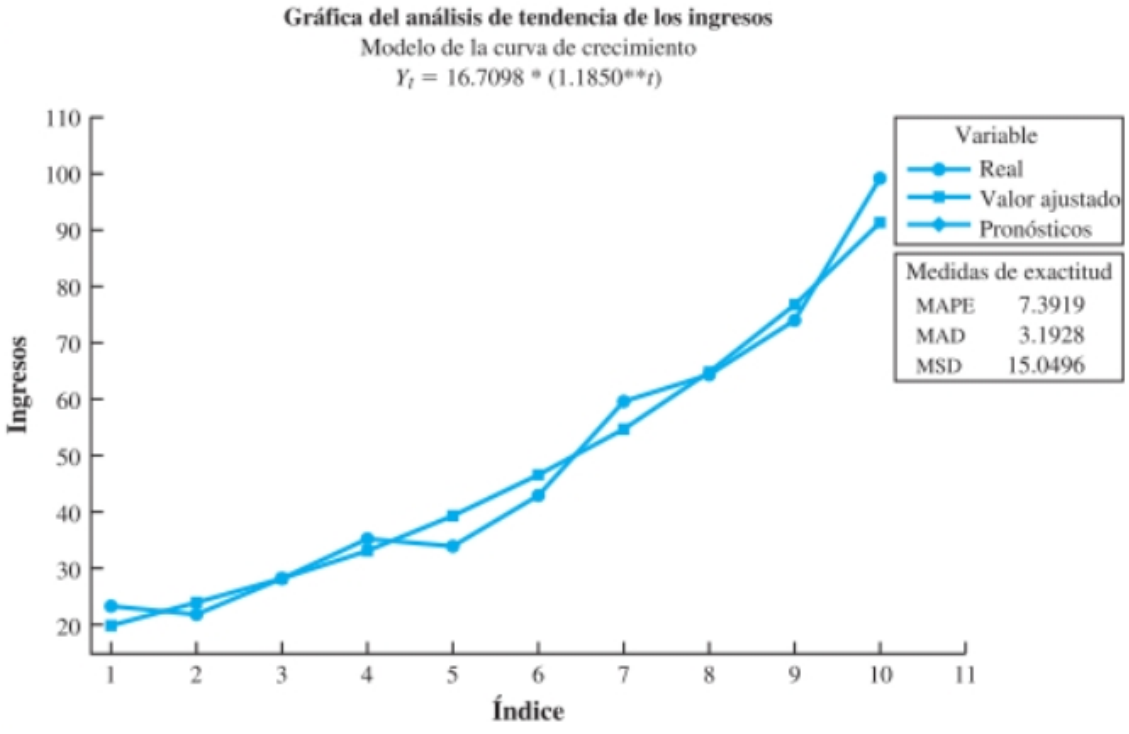
Para entender mejor esta ecuación, suponga que $b_0 = 20$ y $b_1 = 1.2$. Después, para $t = 1$, $T_1 = 20(1.2)^1 = 24$; para $t = 2$, $T_2 = 20(1.2)^2 = 28.8$, y para $t = 3$, $T_3 = 20(1.2)^3 = 34.56$. Observe que T_t no está aumentando una cantidad constante, como en el caso del modelo de tendencia lineal, sino un porcentaje constante; el incremento porcentual es de 20%.

En su módulo de series de tiempo, *MINITAB* tiene la capacidad para calcular una ecuación de tendencia exponencial y puede utilizarse entonces para el pronóstico. Desafortunadamente,

FIGURA 10.12
Resultado de *MINITAB* de la regresión de tendencia cuadrática de la serie de tiempo de ventas de bicicletas

```
The regression equation is
Revenue = 24.2 - 2.11 Year + 0.922 YearSq

Predictor      Coef    SE Coef        T        p
Constant     24.182      4.676     5.17    0.001
Year         -2.106      1.953    -1.08    0.317
YearSq       0.9216     0.1730     5.33    0.001

S = 3.97578   R-Sq = 98.1%   R-Sq(adj) = 97.6%

Analysis of Variance

SOURCE           DF       SS       MS        F       p
Regression        2  5770.1   2885.1   182.52   0.000
Residual Error    7   110.6     15.8
Total             9  5880.8
```

FIGURA 10.13
Resultado del análisis de tendencia de crecimiento exponencial de *MINITAB* para la serie de tiempo de ventas de medicamentos contra el colesterol



Excel no tiene esta capacidad. No obstante, al aplicar logaritmos a los términos de la ecuación (10.8), la metodología del modelo general lineal puede utilizarse para calcular la ecuación de tendencia exponencial.

El módulo de series de tiempo de *MINITAB* es muy fácil de usar para desarrollar una ecuación de tendencia exponencial. No es necesario trabajar con los logaritmos y aplicar el análisis de regresión para calcularla. En la figura 10.13 se muestra la gráfica obtenida por computadora con el procedimiento del análisis de tendencia de series de tiempo de *MINITAB* que se adapta a una ecuación de tendencia exponencial.

10.3 VARIACIÓN ESTACIONAL

WEB file Umbrella

En esta sección se muestra cómo elaborar pronósticos para una serie de tiempo que tiene una variación estacional. En la medida en que existe la estacionalidad, se debe incorporar a los modelos de elaboración de pronósticos para garantizar un pronóstico exacto. Se comenzará por considerar una serie de tiempo estacional sin tendencia y después se estudiará cómo modelar la estacionalidad con la tendencia.

Estacionalidad sin tendencia

Como ejemplo, considere el número de sombrillas vendidas en una tienda de ropa en los últimos cinco años. La tabla 10.12 muestra la serie de tiempo y la figura 10.14 la gráfica correspondiente. Esta última no indica una tendencia a largo plazo en las ventas. De hecho, a menos que se observen cuidadosamente los datos, es posible concluir que éstos siguen una variación horizontal y que podría utilizarse un suavizamiento exponencial sencillo para pronosticar las ventas. Pero una inspección más meticulosa de la gráfica de serie de tiempo revela una variación en los datos. Es decir, el primer y tercer trimestres registran ventas moderadas, el segundo tiene las ventas más altas y el cuarto tiende a ser el trimestre más bajo en volumen de ventas. Por lo tanto, se podría concluir que está presente un patrón estacional trimestral.

Las variables ficticias o binarias (dummy) para tratar las variables cualitativas independientes en un modelo de regresión múltiple pueden utilizar el mismo método para modelar una serie de tiempo con una variación estacional al tratar la estación como variable cualitativa. Recuerde que cuando esta variable tiene k niveles, se necesitan $k = 1$ variables ficticias. Por lo tanto, si hay cuatro estaciones se requieren tres variables ficticias. Por ejemplo, la serie de tiempo de la temporada de ventas de sombrillas es una variable cualitativa con cuatro niveles:

TABLA 10.12 **Serie de tiempo de ventas de sombrillas**

Año	Trimestre	Ventas
1	1	125
	2	153
	3	106
	4	88
2	1	118
	2	161
	3	133
	4	102
3	1	138
	2	144
	3	113
	4	80
4	1	109
	2	137
	3	125
	4	109
5	1	130
	2	165
	3	128
	4	96

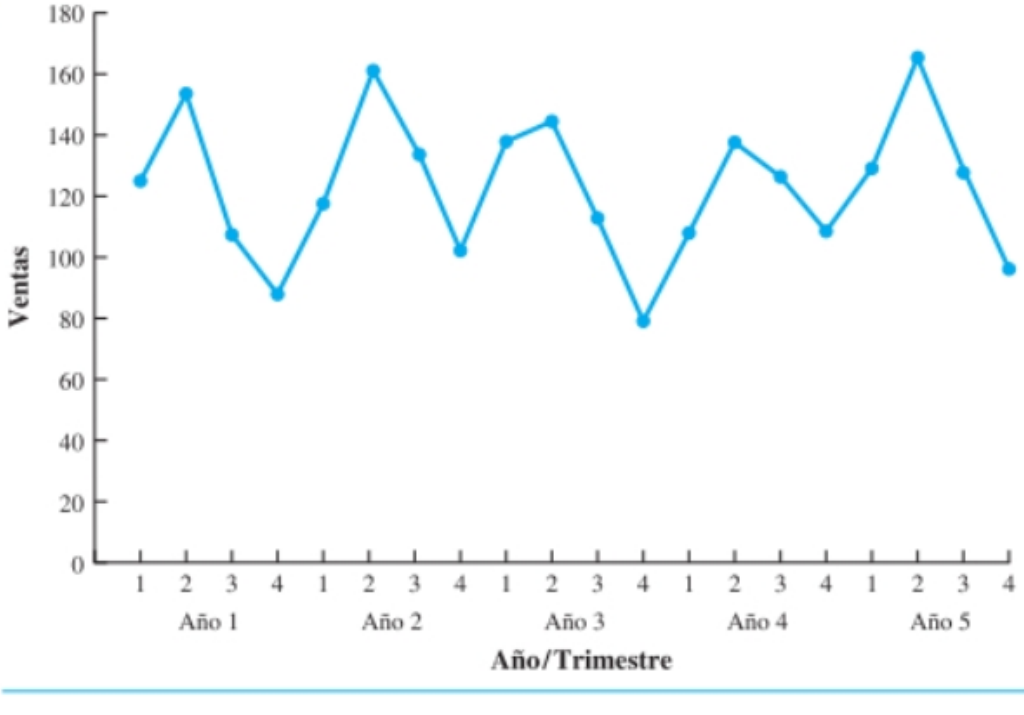
trimestre 1, trimestre 2, trimestre 3 y trimestre 4. Así, para modelar los efectos estacionales en la serie de tiempo de las sombrillas se necesitan $4 - 1 = 3$ variables ficticias. Éstas pueden codificarse de la siguiente manera:

$$\text{Qtr 1} = \begin{cases} 1 \text{ si el trimestre es 1} \\ 0 = \text{en otro caso} \end{cases} \quad \text{Qtr 2} = \begin{cases} 1 \text{ si el trimestre es 2} \\ 0 = \text{en otro caso} \end{cases} \quad \text{Qtr 3} = \begin{cases} 1 \text{ si el trimestre es 3} \\ 0 = \text{en otro caso} \end{cases}$$

Si se utiliza $\hat{Y}$ para denotar el valor estimado o pronosticado de las ventas, la forma general de la ecuación de regresión estimada relacionada con el número de sombrillas vendidas con base en el trimestre Trim(Qtr) en que las ventas ocurren es la siguiente:

$$\hat{Y} = b_0 + b_1 \text{ Qtr1} + b_2 \text{ Qtr2} + b_3 \text{ Qtr3}$$

FIGURA 10.14
Gráfica de la serie de tiempo de ventas de sombrillas



La tabla 10.13 es la serie de tiempo de las ventas de sombrillas que muestra los valores codificados de las variables ficticias. Al utilizar estos datos y el procedimiento de regresión de *MINITAB*, se obtuvo el resultado por computadora que se muestra en la figura 10.15. La ecuación de regresión múltiple estimada obtenida es

$$\text{Ventas} = 95.0 + 29.0\,\text{Qtr1} + 57.0\,\text{Qtr2} + 26.0\,\text{Qtr3}$$

Esta ecuación se puede utilizar para pronosticar las ventas trimestrales para el próximo año.

Trimestre 1. Ventas = 95.0 + 29.0(1) + 57.0(0) + 26.0(0) = 124

Trimestre 2. Ventas = 95.0 + 29.0(0) + 57.0(1) + 26.0(0) = 152

Trimestre 3. Ventas = 95.0 + 29.0(0) + 57.0(0) + 26.0(1) = 121

Trimestre 4. Ventas = 95.0 + 29.0(0) + 57.0(1) + 26.0(0) = 95

Es interesante señalar que se habrían podido obtener los pronósticos trimestrales para el próximo año, simplemente calculando el número promedio de sombrillas que se venden cada trimestre, como se muestra en la siguiente tabla:

Año	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4
1	125	153	106	88
2	118	161	133	102
3	138	144	113	80
4	109	137	125	109
5	130	165	128	96
Promedio	124	152	121	95

Con todo, el resultado de la regresión que se muestra en la figura 10.15 proporciona información adicional que puede utilizarse para evaluar la exactitud del pronóstico y determinar la significancia de los resultados. Y para los tipos más complejos de situaciones, como tratar con una serie de tiempo que tiene efectos tanto de tendencia como estacionales, el método de un promedio simple no funcionará.

TABLA 10.13 **Serie de tiempo con variables ficticias de ventas de sombrillas**

Año	Trimestre	Trim 1	Trim 2	Trim 3	Ventas
1	1	1	0	0	125
	2	0	1	0	153
	3	0	0	1	106
	4	0	0	0	88
2	1	1	0	0	118
	2	0	1	0	161
	3	0	0	1	133
	4	0	0	0	102
3	1	1	0	0	138
	2	0	1	0	144
	3	0	0	1	113
	4	0	0	0	80
4	1	1	0	0	109
	2	0	1	0	137
	3	0	0	1	125
	4	0	0	0	109
5	1	1	0	0	130
	2	0	1	0	165
	3	0	0	1	128
	4	0	0	0	96

FIGURA 10.15
Resultado de la regresión de *MINITAB* para la serie de tiempo de ventas de sombrillas

```
The regression equation is
Sales = 95.0 + 29.0 Qtr1 + 57.0 Qtr2 + 26.0 Qtr3

Predictor     Coef  SE Coef      T      P
Constant    95.000    5.065  18.76  0.000
Qtr1        29.000    7.162   4.05  0.001
Qtr2        57.000    7.162   7.96  0.000
Qtr3        26.000    7.162   3.63  0.002
```

Estacionalidad y tendencia

WEB file TVSales

Ahora ampliaremos el método de regresión para incluir situaciones en las que la serie de tiempo contiene tanto el efecto estacional como una tendencia lineal, y mostraremos cómo pronosticar la serie de tiempo de ventas trimestrales de televisores que se presentó en la sección 10.1. Los datos respectivos se muestran en la tabla 10.14. La gráfica de la serie de tiempo de la figura 10.16 indica que las ventas son muy bajas en el segundo trimestre de cada año y que aumentan en los trimestres 3 y 4. Por lo tanto, se concluye que existe una variación estacional para las ventas de televisores. Pero la serie de tiempo tiene también una tendencia lineal ascendente que tendrá que tomarse en cuenta para formular pronósticos exactos de las ventas trimestrales. Es fácil manejar y combinar el método de las variables ficticias por estacionalidad con el método de regresión de la serie de tiempo para el manejo de la tendencia lineal.

La forma general de la ecuación de regresión múltiple estimada para modelar tanto los efectos estacionales trimestrales como la tendencia lineal en la serie de tiempo de los televisores es la siguiente:

$$\hat{Y}_t = b_0 + b_1\,\text{Qtr1} + b_2\,\text{Qtr2} + b_3\,\text{Qtr3} + b_4 t$$

donde

$\hat{Y}_t$ = estimación o pronóstico de ventas en el periodo t
Qtr1 = 1 si el periodo t corresponde al primer trimestre del año; 0 en caso contrario
Qtr2 = 1 si el periodo t corresponde al segundo trimestre del año; 0 en caso contrario
Qtr3 = 1 si el periodo t corresponde al tercer trimestre del año; 0 en caso contrario
t = periodo

La tabla 10.15 es la serie de tiempo revisada de ventas de televisores que incluye los valores codificados de las variables ficticias y el periodo t. Al utilizar los datos en la tabla 10.15 y el procedimiento de regresión de *MINITAB* se obtuvo el resultado por computadora que se muestra en la figura 10.16. La ecuación de regresión múltiple estimada es

$$\text{Ventas} = 6.07 + 1.36\,\text{Qtr1} + 2.03\,\text{Qtr2} + 0.304\,\text{Qtr3} + 0.146t \qquad (10.9)$$

TABLA 10.14 Serie de tiempo de ventas de televisores

Año	Trimestre	Ventas (miles $)
1	1	4.8
	2	4.1
	3	6.0
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6.0
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8.0
	4	8.4

FIGURA 10.16
Gráfica de la serie de tiempo de ventas de televisores

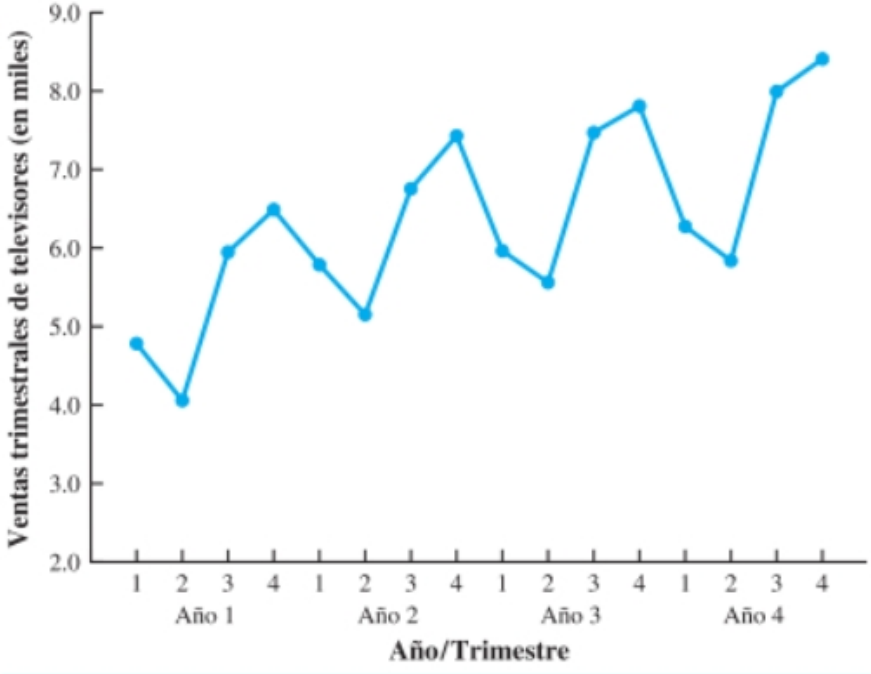


TABLA 10.15 Serie de tiempo con variables ficticias y periodo de ventas de televisores

Año	Trimestre	Trim 1	Trim 2	Trim 3	Periodo	Ventas (miles)
1	1	1	0	0	1	4.8
	2	0	1	0	2	4.1
	3	0	0	1	3	6.0
	4	0	0	0	4	6.5
2	1	1	0	0	5	5.8
	2	0	1	0	6	5.2
	3	0	0	1	7	6.8
	4	0	0	0	8	7.4
3	1	1	0	0	9	6.0
	2	0	1	0	10	5.6
	3	0	0	1	11	7.5
	4	0	0	0	12	7.8
4	1	1	0	0	13	6.3
	2	0	1	0	14	5.9
	3	0	0	1	15	8.0
	4	0	0	0	16	8.4

Ahora se puede utilizar la ecuación (10.9) a efecto de pronosticar las ventas trimestrales para el próximo año, el cual es el año 5 para la serie de tiempo de ventas de televisores; es decir, los periodos 17, 18, 19 y 20.

Pronóstico para el periodo 17 (Trimestre 1 en el año 5)

$$\text{Ventas} = 6.07 - 1.36(1) - 2.03(0) - 0.304(0) + 0.146(17) = 7.19$$

Pronóstico para el periodo 18 (Trimestre 2 en el año 5)

$$\text{Ventas} = 6.07 - 1.36(0) - 2.03(1) - 0.304(0) + 0.146(18) = 6.67$$

FIGURA 10.17
Resultado de la regresión de *MINITAB* de las series de tiempo de ventas de sombrillas

```
The regression equation is
Sales (1000s) = 6.07 - 1.36 Qtr1 - 2.03 Qtr2 - 0.304
                Qtr3 + 0.146 Period

Predictor      Coef  SE Coef       T      P
Constant     6.0688   0.1625   37.35  0.000
Qtr1        -1.3631   0.1575   -8.66  0.000
Qtr2        -2.0337   0.1551  -13.11  0.000
Qtr3        -0.3044   0.1537   -1.98  0.073
Period      0.14562  0.01211   12.02  0.000
```

Pronóstico para el periodo 19 (trimestre 3 en el año 5)

$$\text{Ventas} = 6.07 - 1.36(0) - 2.03(0) - 0.304(1) + 0.146(19) = 8.54$$

Pronóstico para el periodo 20 (trimestre 4 en el año 5)

$$\text{Ventas} = 6.07 - 1.36(0) - 2.03(0) - 0.304(0) + 0.146(20) = 8.99$$

Por lo tanto, si se toma en cuenta los efectos tanto estacionales como de tendencia lineal de las ventas de televisores, las estimaciones de las ventas trimestrales en el año 5 son 7 190, 6 670, 8 540 y 8 990.

Las variables ficticias en la ecuación de regresión múltiple estimada realmente ofrecen cuatro ecuaciones de regresión múltiple estimadas, una para cada trimestre. Por ejemplo, si el periodo t corresponde al trimestre 1, la ecuación estimada para las ventas es

$$\textit{Trimestre}\ 1.\ \text{Ventas} = 6.07 - 1.36(1) - 2.03(0) - 0.304(0) - 0.146t + 4.71 = 0.146t$$

Del mismo modo, si el periodo t corresponde a los trimestres 2, 3 y 4, las estimaciones de las ventas trimestrales son

$$\textit{Trimestre}\ 2.\ \text{Ventas} = 6.07 - 1.36(0) - 2.03(1) - 0.304(0) + 0.146t = 4.04 + 0.146t$$

$$\textit{Trimestre}\ 3.\ \text{Ventas} = 6.07 - 1.36(0) - 2.03(0) - 0.304(1) + 0.146t = 5.77 + 0.146t$$

$$\textit{Trimestre}\ 4.\ \text{Ventas} = 6.07 = 1.36(0) = 2.03(0) = 0.304(0) + 0.146t = 6.07 + 0.146t$$

La pendiente de la tendencia lineal para cada ecuación de pronóstico trimestral es 0.146, lo que indica un crecimiento en las ventas de alrededor de 146 televisores por trimestre. La única diferencia en las cuatro ecuaciones estriba en que tienen diferentes intersecciones. Por ejemplo, la intersección en la ecuación del trimestre 1 es 4.71 y para el trimestre 4 es 6.07. Por lo tanto, las ventas en el trimestre 1 son $4.71 - 6.07 = -1.36$ o 1 360 televisores menos que en el trimestre 4. En otras palabras, el coeficiente estimado en la regresión para Qtr1 en la ecuación (10.9) proporciona una estimación de la diferencia en las ventas entre los trimestres 1 y 4. Interpretaciones similares pueden darse para -2.03, el coeficiente estimado para la variable ficticia Qtr2, y para -0.304, el coeficiente estimado para la variable ficticia Qtr3.

10.4 PATRÓN O COMPONENTE CÍCLICO

En términos matemáticos, el modelo multiplicativo se puede ampliar para incluir el patrón o componente cíclico.

$$Y_t = \text{Tendencia}_t = \text{Cíclica}_t = \text{Estacional}_t = \text{Irregular}_t$$

El componente cíclico, al igual que el estacional, se expresa como un porcentaje de la tendencia. Como se mencionó en la sección 10.1, este componente se debe a los ciclos multianuales en las series de tiempo. Es semejante al componente estacional, pero durante periodos más prolongados. Sin embargo, debido a la extensión del tiempo involucrado, con frecuencia es difícil obtener suficientes datos relevantes para estimar el componente cíclico. Otra dificultad radica en que estos ciclos suelen tener longitudes variables. Como es tan difícil identificar y/o separar los efectos cíclicos de los efectos de tendencia a largo plazo, en la práctica estos efectos a menudo se combinan y se les llama *componente combinado* de tendencia-ciclo.

10.5 VARIACIÓN IRREGULAR

El componente aleatorio mide la variabilidad de las series de tiempo después que se retiran los otros componentes. Contabiliza la variabilidad aleatoria en una serie de tiempo ocasionada por factores imprevistos y no ocurrentes. La mayoría de los componentes irregulares se conforman de variabilidad aleatoria. Sin embargo, ciertos sucesos a veces impredecibles como huelgas, cambios de clima (sequías, inundaciones), terremotos, elecciones o conflictos armados pueden causar irregularidad en una variable.

Movimientos irregulares o al azar o ruido estadístico

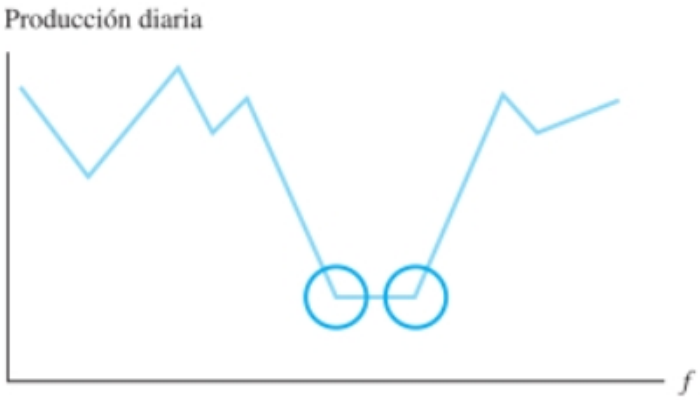
Si bien pueden ser generados por factores de tipo económico, por lo general sus efectos producen variaciones que sólo duran un corto intervalo de tiempo. Claro que debe reconocerse que en ocasiones sus efectos sobre el comportamiento de una serie pueden ser tan intensos que podrían dar lugar con cierta facilidad a un nuevo ciclo o a otros movimientos.

Por tanto, al analizar una serie de tiempo es necesario tener en cuenta el comportamiento de cada uno de estos componentes, para lo cual el criterio más lógico a seguir es aislarlos de manera secuencial partiendo de la serie original para luego analizarlos de manera individual. Si bien esto supone la utilización de métodos estadísticos adecuados, la mejor forma de apreciarlos es a través de la observación visual.

La detección outlier se refiere a puntos de una serie que se escapan de lo normal. Un outlier es una observación de la serie que corresponde a un comportamiento anormal del fenómeno (sin incidencias futuras) o a un error de medición.

Se debe determinar desde fuera si un punto dado es outlier o no. Si se concluye que lo es, debe omitirse o reemplazarse por otro valor antes de analizar la serie. Por ejemplo, en un estudio de la producción diaria en una fábrica se presentó una situación que dio origen a la gráfica de la figura 10.18.

FIGURA 10.18
Análisis de producción en una fábrica.



Los dos puntos destacados con un círculo parecen corresponder a un comportamiento anormal de la serie. Sin embargo, al investigar ambos puntos se pudo comprobar que correspondían a dos días de paro, lo cual afectó por tanto la producción en esos días. El problema fue solucionado al eliminar las observaciones e interpolar.

10.6 PROBLEMAS QUE INCLUYEN LOS CUATRO COMPONENTES

Esta sección inicia con la elaboración de pronósticos de la serie de tiempo de ventas de gasolina que se muestra en la tabla 10.1, utilizando el más simple de todos los métodos de elaboración de pronósticos: uno que utiliza el volumen de ventas de la última semana como predictor de la siguiente semana. Por ejemplo, un distribuidor vendió 17 mil galones de gasolina en la semana 1; este valor se utiliza como el pronóstico para la semana 2. Después, se utiliza 21, el valor real de las ventas en la semana 2, como el pronóstico para la semana 3, y así sucesivamente. Los pronósticos obtenidos con este método para los datos históricos se muestran en la columna Pronóstico en la tabla 10.16. Debido a su sencillez, a este enfoque se le denomina a menudo método de elaboración de pronósticos *ingenuo*.

¿Qué tan exactos son los pronósticos al utilizar el método *ingenuo*? Para responder esta pregunta se presentan varias medidas de exactitud de los pronósticos. Estas medidas se utilizan para determinar qué tan bien un método puede reproducir los datos de las series de tiempo disponibles. Al seleccionar el método con la mejor exactitud de los datos ya conocidos, se espera que aumente la probabilidad de obtener un mejor pronóstico para periodos futuros.

El concepto clave relacionado con la medida de exactitud del pronóstico es el **error de pronóstico**, que se define como

$$\text{Error de pronóstico} = \text{Valor real} - \text{Pronóstico}$$

Por ejemplo, debido a que el distribuidor vendió en realidad 21 mil galones de gasolina en la semana 2 y el pronóstico al utilizar el volumen de ventas en la semana 1 fue de 17 mil galones, el error de pronóstico en la semana 2 es

$$\text{Error de pronóstico en la semana 2} = 21 - 17 = 4$$

El hecho de que el error de pronóstico sea positivo, indica que en la semana 2 el método de elaboración de pronósticos subestimó el valor real de las ventas. A continuación utilice 21, el valor real de las ventas en la semana 2, como pronóstico para la semana 3. Ya que el valor real de las ventas en la semana 3 es 19, el error de pronóstico para esta semana es $19 - 21 = -2$. En este caso, el error negativo indica que en la semana 3 el pronóstico sobrestimó el valor real. Así, el error de pronóstico puede ser positivo o negativo dependiendo de si es demasiado bajo o demasiado alto. Un resumen completo de los errores de pronóstico para este método ingenuo se muestra en la tabla 10.16, en la columna Error de pronóstico.

TIP MI CONSEJO

En el análisis de regresión, un residual se define como la diferencia entre el valor observado y el valor estimado de la variable dependiente. Los errores de pronóstico son análogos a los residuales en el análisis de regresión.

Una medida sencilla de exactitud de los pronósticos es la media o promedio de errores de pronóstico. La tabla 10.16 muestra que la suma de estos errores para la serie de tiempo de las ventas de gasolina es 5, por lo que la media o promedio del error de pronóstico es $5/11 = 0.45$. Observe que aunque la serie de tiempo de gasolina se compone de 12 valores, al calcular la media del error se divide entre 11 la suma de los errores, ya que existen sólo 11 errores de pronóstico. Debido a que la media del error de pronóstico es positiva, el método genera pronósticos bajos; es decir, los valores observados tienden a ser mayores que los pronosticados. Dado que los errores de pronóstico positivos y negativos tienden a anularse entre sí, es probable que la media del error sea pequeña, así que ésta no es una medida muy útil de la exactitud del pronóstico.

El **error absoluto medio**, que se denota EAM, es una medida de exactitud del pronóstico que evita el problema de los errores positivos y negativos que se anulen entre sí. Como es de esperar, dado su nombre, EAM es el promedio de los valores absolutos de los errores de pronóstico. La tabla 10.16 muestra que la suma de los valores absolutos de los errores de pronóstico es 41; por lo tanto

$$\text{EAM} = \text{promedio del valor absoluto de los errores de pronóstico} = \frac{41}{11} = 3.73$$

TABLA 10.16 **Cálculos y medidas de exactitud de pronósticos utilizando el valor más reciente como pronóstico para el siguiente periodo**

Semana	Valor de la serie de tiempo	Pronóstico	Error de pronóstico	Valor absoluto del error de pronóstico	Error cuadrado de pronóstico	Error porcentual	Valor absoluto del error porcentual
1	17						
2	21	17	4	4	16	19.05	19.05
3	19	21	−2	2	4	−10.53	10.53
4	23	19	4	4	16	17.39	17.39
5	18	23	−5	5	25	−27.78	27.78
6	16	18	−2	2	4	−12.50	12.50
7	20	16	4	4	16	20.00	20.00
8	18	20	−2	2	4	−11.11	11.11
9	22	18	4	4	16	18.18	18.18
10	20	22	−2	2	4	−10.00	10.00
11	15	20	−5	5	25	−33.33	33.33
12	22	15	7	7	49	31.82	31.82
		Totales	5	41	179	1.19	211.69

TIP MI CONSEJO

En el análisis de regresión, el cuadrado medio debido al error (CME) o error cuadrático medio, es la suma de los residuales cuadrados dividida entre sus grados de libertad. En el pronóstico, el CME es el promedio de la suma de los errores cuadrados de pronóstico.

Otra medida que evita el problema de los errores de pronóstico positivos y negativos que se compensan entre sí se obtiene al calcular el promedio de los errores cuadrados de pronóstico. Esta medida de exactitud en los pronósticos se llama **cuadrado medio debido al error**, que se denota como CME, o error cuadrático medio. En la tabla 10.16, la suma de los cuadrados debido al error es 179, por lo tanto,

$$\text{CME} = \text{promedio de la suma de los errores cuadrados de pronóstico} = \frac{179}{11} = 16.27$$

El tamaño del EAM y del CME depende de la escala de los datos. Como resultado, es difícil realizar comparaciones de los distintos intervalos de tiempo, como la de un método de elaboración de pronósticos de ventas mensuales de gasolina con un método de elaboración de pronósticos de ventas semanales, o realizar comparaciones de las distintas series de tiempo. Para realizar comparaciones como éstas se debe trabajar con las medidas relativas o porcentuales de los errores. El **error porcentual absoluto medio**, que se denota EPAM, es una medida de este estilo. Para calcular el EPAM, en primer lugar se debe determinar el error porcentual de cada pronóstico. Por ejemplo, el error porcentual que corresponde al pronóstico de 17 en la semana 2 se calcula dividiendo el error de pronóstico en la semana 2 entre el valor real en la semana 2 y multiplicando el resultado por 100. Para esta semana, el error porcentual se calcula de la siguiente manera:

$$\text{Error porcentual para la semana 2} = \frac{4}{21}(100) = 19.05\%$$

Así, el error de pronóstico para la semana 2 es 19.05% del valor observado en tal semana. Un resumen completo de los errores porcentuales se muestra en la tabla 10.16, en la columna Error porcentual. En la columna que le sigue se muestra el valor absoluto de este error porcentual.

La tabla 10.16 indica que la suma de los valores absolutos de los errores porcentuales es 211.69, por lo que

$$\text{EPAM} = \text{promedio del valor absoluto de los errores porcentuales de pronóstico} = \frac{211.69}{11} = 19.24\%$$

En resumen, al utilizar el método de elaboración de pronósticos ingenuo (la más reciente observación), se obtuvieron las siguientes medidas de exactitud del pronóstico:

$$\text{EAM} = 3.73$$
$$\text{CME} = 16.27$$
$$\text{EPAM} = 19.24\%$$

Estas medidas de exactitud miden simplemente qué tan bien el método de elaboración de pronósticos puede pronosticar los valores históricos de las series de tiempo. Ahora, suponga que se desea pronosticar las ventas para un periodo futuro, como la semana 13. En este caso, el pronóstico es 22, el valor real de las series de tiempo en la semana 12. ¿Es ésta una estimación exacta de las ventas para la semana 13? Desafortunadamente no hay manera de abordar el tema de la exactitud relacionada con el pronóstico para periodos futuros. Pero si se elige un método de elaboración de pronósticos que funcione bien para los datos históricos, y se piensa que el patrón histórico continuará en el futuro, se deben obtener resultados que, en última instancia, probaron ser buenos.

Antes de concluir esta sección se considerará otro método de elaboración de pronósticos para las series de tiempo de ventas de gasolina de la tabla 10.1 Suponga que se utiliza el promedio de todos los datos históricos disponibles como pronóstico para el próximo periodo. Comience por elaborar un pronóstico para la semana 2. Ya que existe sólo un valor histórico disponible antes de la semana 2, el pronóstico para ésta es sólo el valor de la serie de tiempo para la semana 1, por lo que el pronóstico es de 17 mil galones de gasolina. Para calcular el pronóstico de la semana 3 se toma el promedio de los valores de ventas de las semanas 1 y 2. Por lo tanto, el resultado que se obtiene es el que se indica después de la tabla.

TABLA 10.17 **Cálculo y medidas de exactitud del pronóstico al utilizar el promedio de todos los datos históricos como pronóstico para el siguiente periodo**

Semana	Valor de la serie de tiempo	Pronóstico	Error de pronóstico	Valor absoluto del error de pronóstico	Error cuadrado de pronóstico	Error porcentual	Valor absoluto del error porcentual
1	17						
2	21	17.00	4.00	4.00	16.00	19.05	19.05
3	19	19.00	0.00	0.00	0.00	0.00	0.00
4	23	19.00	4.00	4.00	16.00	17.39	17.39
5	18	20.00	−2.00	2.00	4.00	−11.11	11.11
6	16	19.60	−3.60	3.60	12.96	−22.50	22.50
7	20	19.00	1.00	1.00	1.00	5.00	5.00
8	18	19.14	−1.14	1.14	1.31	−6.35	6.35
9	22	19.00	3.00	3.00	9.00	13.64	13.64
10	20	19.33	0.67	0.67	0.44	3.33	3.33
11	15	19.40	−4.40	4.40	19.36	−29.33	29.33
12	22	19.00	3.00	3.00	9.00	13.64	13.64
		Totales	4.53	26.81	89.07	2.76	141.34

$$\text{Pronóstico para la semana } 3 = \frac{17 + 21}{2} = 19$$

De la misma forma, el pronóstico para la semana 4 es

$$\text{Pronóstico para la semana } 4 = \frac{17 + 21 + 19}{3} = 19$$

Los pronósticos que se obtienen al utilizar este método para las series de tiempo de ventas de gasolina se muestran en la tabla 10.17 en la columna Pronóstico. Con estos resultados se obtuvieron los siguientes valores de EAM, CME y EPAM.

$$\text{EAM} = \frac{26.81}{11} = 2.44$$

$$\text{CME} = \frac{89.07}{11} = 8.10$$

$$\text{EPAM} = \frac{141.34}{11} = 12.85\%$$

Ahora se puede determinar la exactitud de los dos métodos de elaboración de pronósticos que se han considerado en esta sección mediante la comparación de los valores de EAM, CME y EPAM.

	Método ingenuo	Promedio de los valores pasados
EAM	3.73	2.44
CME	16.27	8.10
EPAM	19.24%	12.85%

Para cada medida, el promedio de los valores pasados proporciona pronósticos más precisos que al utilizar la observación más reciente como pronóstico para el próximo periodo. En general, si la serie de tiempo subyacente es estacionaria, el promedio de todos los datos históricos siempre proporcionará los mejores resultados.

Pero suponga que la serie de tiempo subyacente no es estacionaria. En la sección 10.1 se menciona que las variaciones en las condiciones de negocios suelen dar lugar a una serie de tiempo con una variación horizontal que cambia a un nuevo nivel. Se estudió una situación en la que el distribuidor de gasolina firmó un contrato con la policía del estado de Vermont para suministrar combustible a las patrullas de la policía del sur del estado. La tabla 10.2 muestra

el número de galones de gasolina que se vendieron para la serie de tiempo original y para las 10 semanas después de firmado el nuevo contrato, y la figura 10.2 presenta la gráfica que corresponde a las series de tiempo. Observe el cambio en el nivel de la semana 13 para la serie de tiempo resultante. Cuando ocurre este cambio, le toma tiempo al método de elaboración de pronósticos que utiliza el promedio de todos los datos históricos ajustarse a un nuevo nivel de series de tiempo. Pero en este caso, el método ingenuo simple se ajusta muy rápidamente a los cambios en el nivel, debido a que utiliza como pronóstico la observación más reciente.

Las medidas de exactitud de los pronósticos son factores importantes en la comparación de distintos métodos de elaboración de los mismos, pero se debe tener cuidado de no depender demasiado de ellas. El buen juicio y el conocimiento sobre las condiciones de negocios que puedan afectar el pronóstico también deben tomarse en cuenta cuidadosamente cuando se elija un método. La exactitud de los pronósticos históricos no es la única consideración, sobre todo si es probable que la serie de tiempo cambie en el futuro.

Existen métodos más sofisticados para elaborar los pronósticos de una serie de tiempo que muestren un patrón horizontal. Algunos de ellos tienen la ventaja de adaptarse a situaciones donde las series de tiempo cambian a un nuevo nivel. La capacidad de un método de pronósticos para adaptarse rápidamente a estos cambios es una consideración importante, especialmente en situaciones de elaboración de pronósticos a corto plazo.

10.7 ANÁLISIS DE LAS SERIES DE TIEMPO EN LOS PRONÓSTICOS

En esta sección se enfoca la atención en lo que se conoce como **descomposición de series de tiempo**, la cual se puede utilizar para separar o descomponer una serie de tiempo en sus componentes de tendencia, estacional e irregular; aunque este método puede utilizarse para el pronóstico, su aplicación principal es obtener una mejor comprensión de la serie de tiempo. Muchas series de tiempo para las empresas y la economía de Estados Unidos son mantenidas y publicadas por agencias gubernamentales, como la Oficina del Censo y la Oficina de Estadísticas Laborales. Estas agencias utilizan la descomposición de series de tiempo para crear series desestacionalizadas.

Entender qué sucede en realidad con una serie de tiempo a menudo depende del uso de los datos desestacionalizados. Por ejemplo, podríamos estar interesados en saber si el consumo de energía eléctrica está aumentando en nuestra área. Suponga que se entera de que éste se redujo 3% en septiembre respecto al mes anterior. Se debe ser cuidadoso al ejercitar el uso de esa información, porque cada vez que se presenta una influencia estacional, esas comparaciones pueden ser engañosas si los datos no han sido desestacionalizados. El hecho de que el consumo de energía eléctrica se haya reducido 3% entre agosto y septiembre podría deberse al efecto estacional que se relaciona con una disminución del uso del aire acondicionado y no a una disminución del uso de la energía eléctrica a largo plazo. En efecto, después de ajustar el efecto estacional, se podría incluso determinar que aumentó el consumo de energía eléctrica. Muchas series de tiempo, como las estadísticas de desempleo y las ventas de casas y al detalle (o al menudeo), están sujetas a fuertes influencias estacionales. Es importante desestacionalizar dichos datos, antes de emitir un juicio acerca de cualquier tendencia a largo plazo.

Los métodos de descomposición de series tiempo asumen que Y_t, el valor real de la serie de tiempo en el periodo t, es una función de tres componentes: un componente de tendencia; un componente estacional, y un componente irregular o de error. El cómo se combinan estos tres componentes para generar los valores observados de la serie de tiempo depende de si se asume que la relación entre ellos se describe mejor por un modelo aditivo o un modelo multiplicativo. Un **modelo de descomposición aditivo** tiene la siguiente forma:

$$Y_t = \text{Tendencia}_t = \text{Estacional}_t = \text{Irregular}_t \qquad (10.11)$$

donde

$$\begin{aligned}\text{Tendencia}_t &= \text{valor de la tendencia en el periodo } t\\ \text{Estacional}_t &= \text{valor estacional en el periodo } t\\ \text{Irregular}_t &= \text{valor irregular en el periodo } t\end{aligned}$$

TIP MI CONSEJO

El componente irregular corresponde al término de error α en el modelo de regresión lineal simple.

En el modelo aditivo, los valores de los tres componentes simplemente se suman para obtener el valor real de la serie de tiempo Y_t. El componente irregular o de error toma en cuenta la variabilidad de la serie de tiempo que no puede explicarse por los componentes de tendencia y estacional.

Un modelo aditivo es apropiado en situaciones en las que las fluctuaciones estacionales no dependen del nivel de la serie de tiempo. El modelo de regresión que incorpora efectos estacionales y de tendencia en la sección 10.3 es un modelo aditivo. Éste es adecuado si las fluctuaciones estacionales en el periodo anterior son casi del mismo tamaño que las fluctuaciones estacionales en periodos posteriores. Sin embargo, si las fluctuaciones estacionales cambian en el tiempo y son cada vez mayores a medida que aumenta el volumen de ventas debido a una tendencia lineal a largo plazo, entonces se debe utilizar el modelo multiplicativo. Muchas series de tiempo para las empresas y la economía siguen esta variación.

Un **modelo de descomposición multiplicativo** toma la siguiente forma:

$$Y_t = \text{Tendencia}_t \times \text{Estacional}_t \times \text{Irregular}_t \qquad (10.12)$$

donde

$$\begin{aligned}\text{Tendencia}_t &= \text{valor de la tendencia en el periodo } t\\ \text{Estacional}_t &= \text{índice estacional en el periodo } t\\ \text{Irregular}_t &= \text{índice irregular en el periodo } t\end{aligned}$$

TIP MI CONSEJO

La Oficina del Censo utiliza un modelo multiplicativo, en conjunción con su metodología para desestacionalizar las series de tiempo.

En este modelo los componentes de tendencia, estacional e irregular se multiplican para dar el valor de la serie de tiempo. La tendencia se mide en las unidades de producto de la serie que se pronostica. Sin embargo, los componentes estacional e irregular se miden en términos relativos, con valores superiores a 1.00 que indican los efectos por encima de la tendencia y valores menores a 1.00 que indican los efectos por debajo de la tendencia.

Debido a que éste es el método que más se utiliza en la práctica, nuestro análisis de descomposición de las series de tiempo se limitará a mostrar cómo se desarrollan las estimaciones de los componentes de tendencia y estacional de un modelo multiplicativo. A modo de ejemplo, se trabajará con la serie de tiempo de las ventas trimestrales de televisores que se presentó en la sección 10.3; los datos de las ventas trimestrales se muestran en la tabla 10.14, y la gráfica correspondiente de la serie de tiempo en la figura 10.16. Después de demostrar cómo se descompone una serie de tiempo con el modelo multiplicativo, se estudiará cómo pueden combinarse los índices estacionales y el componente de tendencia para elaborar un pronóstico.

Cálculo de los índices estacionales

La figura 10.16 indica que las ventas son menores en el segundo trimestre de cada año y aumentan en los trimestres 3 y 4. Por lo tanto, se concluye que existe una variación estacional para la serie de tiempo de las ventas de televisores. El procedimiento que se utiliza para determinar la influencia estacional de cada trimestre comienza por calcular un promedio móvil para separar los componentes estacional e irregular de los datos, lo que deja una serie de tiempo que contiene sólo la tendencia y cualquier variación aleatoria restante que no fue eliminada por los cálculos del promedio móvil.

Como se trabaja con series trimestrales, se utilizarán cuatro valores en cada promedio móvil. El cálculo del promedio móvil de los primeros cuatro trimestres de ventas de televisores es

$$\text{Primer promedio móvil} = \frac{4.8 + 4.1 + 6.0 + 6.5}{4} = \frac{21.4}{4} = 5.35$$

Observe que el cálculo del promedio móvil de los primeros cuatro trimestres da el promedio trimestral de las ventas durante el año 1 de la serie de tiempo. Para continuar con este cálculo se agrega el valor 5.8 correspondiente al primer trimestre del año 2 y se elimina el 4.8 del primer trimestre del año 1. Por lo tanto, el segundo promedio móvil es

$$\text{Segundo promedio móvil} = \frac{4.1 + 6.0 + 6.5 + 5.8}{4} = \frac{22.4}{4} = 5.60$$

De modo similar, el cálculo del tercer promedio móvil es (6.0 = 6.5 = 5.8 = 5.2)/4 = 5.875.

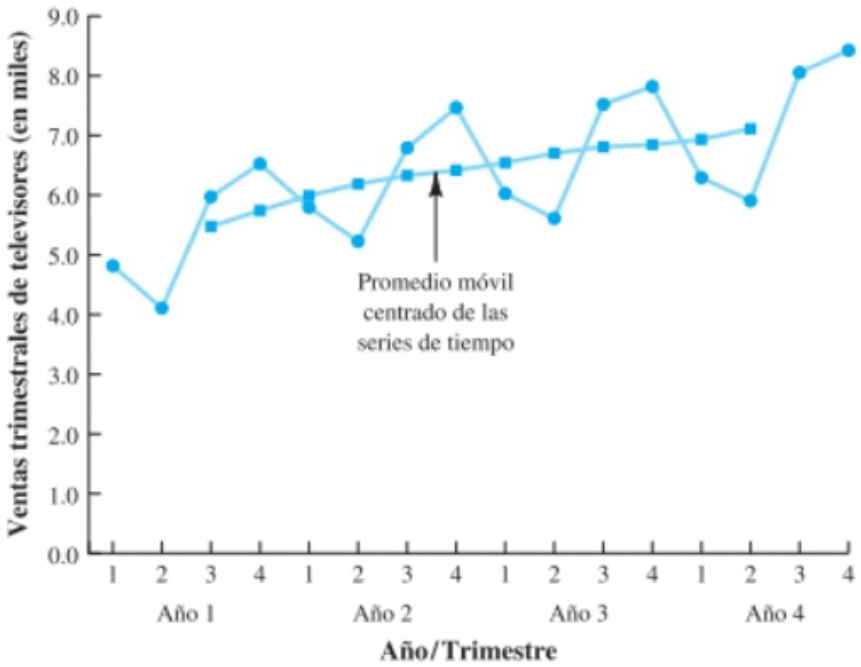
Antes de continuar con el cálculo de los promedios móviles de toda la serie de tiempo, vuelva al primero que resultó en un valor de 5.35. Éste es el promedio trimestral del volumen de ventas para el año 1. Al retroceder en su cálculo, parece razonable asociar el valor 5.35 con el "central" del grupo del promedio móvil. Sin embargo, observe que como en cada promedio móvil intervienen cuatro trimestres, no hay trimestre central. El valor 5.35 corresponde en realidad al periodo 2.5, la segunda mitad del trimestre 2 y la primera mitad del trimestre 3. Asimismo, al pasar al valor del siguiente promedio móvil, que es 5.60, el trimestre central corresponderá al periodo 3.5, la última mitad del trimestre 3 y la primera mitad del 4.

Los dos valores del promedio móvil que se calculan no corresponden directamente a los trimestres originales de la serie de tiempo. Esta dificultad se resuelve calculando el promedio de los dos promedios móviles. Ya que el centro del primero es el periodo 2.5 (la mitad de un periodo o trimestre más temprano) y el centro del segundo es el periodo 3.5 (la mitad del periodo o trimestre más tarde), el promedio de los dos promedios móviles se centra en el trimestre 3, exactamente donde debe estar. Este promedio se conoce como *promedio móvil centrado*, y para el periodo 3 es (5.35 + 5.60)/2 = 5.475, mientras que para el periodo 4 es (5.60 + 5.875)/2 = 5.738. La tabla 10.18 muestra un resumen completo de los cálculos del promedio móvil y del promedio móvil centrado para los datos de las ventas de televisores.

TABLA 10.18 **Cálculos de los promedios móviles centrados de la serie de tiempo de ventas de televisores**

Año	Trimestre	Ventas (en miles)	Promedio móvil de cuatro trimestres	Promedio móvil centrado
1	1	4.8		
1	2	4.1		
			5.350	
1	3	6.0		5.475
			5.600	
1	4	6.5		5.738
			5.875	
2	1	5.8		5.975
			6.075	
2	2	5.2		6.188
			6.300	
2	3	6.8		6.325
			6.350	
2	4	7.4		6.400
			6.450	
3	1	6.0		6.538
			6.625	
3	2	5.6		6.675
			6.725	
3	3	7.5		6.763
			6.800	
3	4	7.8		6.838
			6.875	
4	1	6.3		6.938
			7.000	
4	2	5.9		7.075
			7.150	
4	3	8.0		
4	4	8.4		

FIGURA 10.19
Series de tiempo de ventas trimestrales de televisores y su promedio móvil centrado



¿Qué información se obtiene de los promedios móviles centrados de la tabla 10.18 de esta serie de tiempo? La figura 10.19 muestra una gráfica de los valores reales de las series de tiempo y de los valores de los promedios móviles centrados. Observe, sobre todo, cómo estos últimos tienden a "suavizar" tanto las fluctuaciones estacionales como las irregulares de la serie de tiempo. Los promedios móviles centrados representan la tendencia en los datos y cualquier variación aleatoria que no se ha eliminado con el uso de los promedios móviles para suavizar los datos.

Antes se demostró que el modelo de descomposición multiplicativa es

$$Y_t = \text{Tendencia}_t \times \text{Estacional}_t \times \text{Irreglar}_t$$

TIP MI CONSEJO

Los valores del componente irregular estacional son llamados a menudo valores sin tendencia de las series de tiempo.

Al dividir cada lado de esta ecuación entre el componente de tendencia T_t, se puede identificar el efecto estacional irregular en las series de tiempo.

$$\frac{Y_t}{\text{Tendencia}_t} = \frac{\text{Tendencia}_t \times \text{Estacional}_t \times \text{Irregular}_t}{\text{Tendencia}_t} = \text{Estacional}_t \times \text{Irregular}_t$$

Por ejemplo, el tercer trimestre del año 1 muestra un valor de tendencia de 5.475 (el promedio móvil centrado). Así 6.0/5.475 = 1.096 es el valor combinado estacional-irregular del componente irregular. En la tabla 10.19 se resumen los valores del componente estacional-irregular de toda la serie de tiempo.

Considere los valores del componente estacional-irregular para el tercer trimestre: 1.096, 1.075 y 1.109. Los valores de la parte estacional-irregular mayores que 1.00 indican efectos por encima de la tendencia estimada, y los valores menores que 1.00 indican efectos por debajo de la tendencia estimada. Así, los tres valores del componente estacional-irregular para el trimestre 3 muestran un efecto por encima del promedio en el tercer trimestre. Ya que cada año las fluctuaciones en los valores estacional-irregular se deben principalmente al error aleatorio, se pueden promediar los valores calculados para eliminar la influencia irregular y obtener una estimación de la influencia estacional del tercer trimestre.

$$\text{Efecto estacional del trimestre 3} = \frac{1.096 + 1.075 + 1.109}{3} = 1.09$$

Al número 1.09 se le conoce como *índice estacional* del tercer trimestre. La tabla 10.20 resume los cálculos necesarios para obtener los índices estacionales de la serie de tiempo de ventas de televisores. Los índices estacionales de los cuatro trimestres son 0.93, 0.84, 1.09 y 1.14.

TABLA 10.19

Valores del componente estacional-irregular de la serie de tiempo de ventas de televisores

Año	Trimestre	Ventas (en miles)	Promedio móvil centrado	Valor estacional-irregular
1	1	4.8		
1	2	4.1		
1	3	6.0	5.475	1.096
1	4	6.5	5.738	1.133
2	1	5.8	5.975	0.971
2	2	5.2	6.188	0.840
2	3	6.8	6.325	1.075
2	4	7.4	6.400	1.156
3	1	6.0	6.538	0.918
3	2	5.6	6.675	0.839
3	3	7.5	6.763	1.109
3	4	7.8	6.838	1.141
4	1	6.3	6.938	0.908
4	2	5.9	7.075	0.834
4	3	8.0		
4	4	8.4		

La interpretación de los índices estacionales en la tabla ofrece una idea sobre el componente estacional de las ventas de televisores. El mejor trimestre de ventas es el cuarto, con ventas promedio de 14% por encima de la tendencia estimada. El peor, o más bajo, es el segundo trimestre; su índice estacional de 0.84 indica que el promedio de ventas se encuentra 16% por debajo de la tendencia estimada. El componente estacional se corresponde claramente con la expectativa intuitiva de que el interés por ver televisión y, por tanto, los patrones de compra de televisores tienden a alcanzar el punto máximo en el cuarto trimestre, debido a la próxima temporada de invierno y a la reducción de las actividades al aire libre. Las bajas ventas del segundo trimestre reflejan un menor interés por ver televisión dadas las actividades de primavera y antes del verano de los clientes potenciales.

Algunas veces es necesario un último ajuste para obtener los índices estacionales. Debido a que el modelo multiplicativo requiere que el índice estacional promedio sea igual a 1.00, la suma de los cuatro índices de la tabla 10.20 debe ser igual a 4.00. En otras palabras, los efectos estacionales deben incluso nivelarse a lo largo del año. En el ejemplo visto aquí, el promedio de los índices estacionales es igual a 1.00, y por lo tanto no se requiere ningún tipo de ajuste. En otros casos puede requerirse un ligero ajuste. Para realizarlo, se multiplica cada índice estacional por el número de estaciones, dividido entre la suma de los índices estacionales sin ajustar. Por ejemplo, cuando se tienen datos trimestrales, se multiplica cada índice estacional por 4/(suma de los índices estacionales no ajustados). En algunos ejercicios se requerirá hacer este ajuste para obtener el índice estacional adecuado.

Desestacionalización de una serie de tiempo

Una serie de tiempo a la que se le han eliminado los efectos estacionales se conoce como **serie de tiempo desestacionalizada**, y al proceso de uso de los índices estacionales para eliminar los efectos estacionales de una serie de tiempo se le conoce como *desestacionalizar la serie de tiempo*. Al utilizar un modelo de descomposición multiplicativa se desestacionaliza una serie de

TABLA 10.20

Cálculo de los índices estacionales de la serie de tiempo de ventas de televisores

Trimestre	Valores estacional-irregular			Índice estacional
1	0.971	0.918	0.908	0.93
2	0.840	0.839	0.834	0.84
3	1.096	1.075	1.109	1.09
4	1.133	1.156	1.141	1.14

TABLA 10.21

Valores desestacionalizados de las series de tiempo de ventas de televisores

Año	Trimestre	Periodo	Ventas (en miles)	Índice estacional	Ventas desestacionalizadas
1	1	1	4.8	0.93	5.16
	2	2	4.1	0.84	4.88
	3	3	6.0	1.09	5.50
	4	4	6.5	1.14	5.70
2	1	5	5.8	0.93	6.24
	2	6	5.2	0.84	6.19
	3	7	6.8	1.09	6.24
	4	8	7.4	1.14	6.49
3	1	9	6.0	0.93	6.45
	2	10	5.6	0.84	6.67
	3	11	7.5	1.09	6.88
	4	12	7.8	1.14	6.84
4	1	13	6.3	0.93	6.77
	2	14	5.9	0.84	7.02
	3	15	8.0	1.09	7.34
	4	16	8.4	1.14	7.37

TIP MI CONSEJO

Las series de tiempo económicas ajustadas por variaciones estacionales son generalmente reportadas en publicaciones como el *Survey of Current Business, The Wall Street Journal* y *BusinessWeek*.

tiempo dividiendo cada observación entre el índice estacional correspondiente. El modelo de descomposición multiplicativo es

$$Y_t = \text{Tendencia}_t = \text{Estacional}_t = \text{Irregular}_t$$

Así, cuando se divide cada una de las observaciones de la serie de tiempo (Yt) entre su índice estacional correspondiente, los datos resultantes muestran sólo la tendencia y la variabilidad aleatoria (el componente irregular). Las series de tiempo desestacionalizadas de ventas de televisores se resumen en la tabla 10.21. Una gráfica de las series de tiempo desestacionalizadas se muestra en la figura 10.20.

Uso de una serie de tiempo desestacionalizada para identificar tendencias

La gráfica de series de tiempo desestacionalizadas de ventas de televisores que se muestra en la figura 10.20 parece tener una tendencia lineal ascendente. Para identificar esta tendencia, se ajustará una ecuación de tendencia lineal de las series de tiempo desestacionalizadas utilizan-

FIGURA 10.20

Series de tiempo desestacionalizadas de ventas de televisores

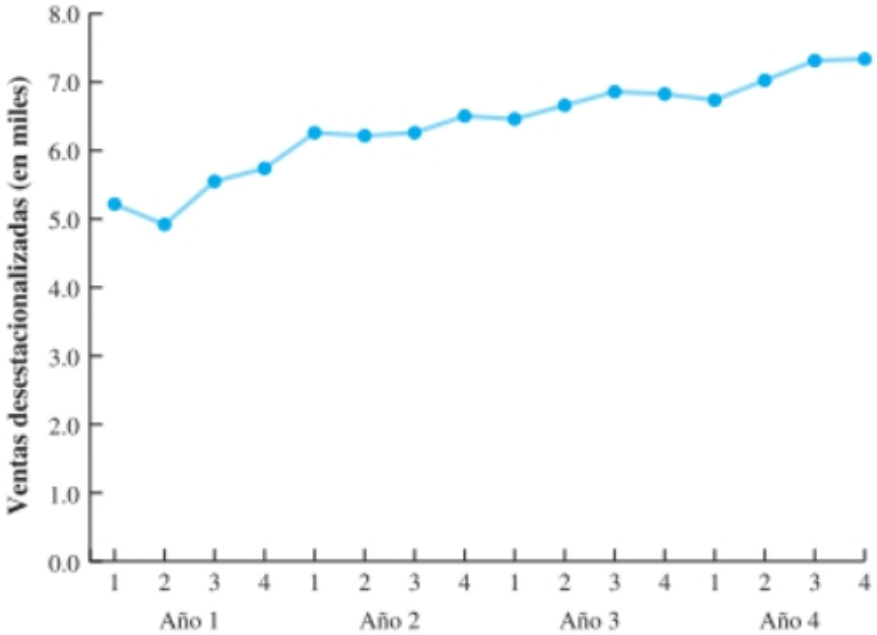


FIGURA 10.21
Resultados de la regresión de *MINITAB* de la serie de tiempo desestacionalizada de ventas de televisores

```
The regression equation is
Deseasonalized Sales = 5.10 + 0.148 Period

Predictor      Coef  SE Coef      T      P
Constant     5.1050   0.1133  45.07  0.000
Period      0.14760  0.01171  12.60  0.000

S = 0.215985   R-Sq = 91.9%   R-Sq(adj) = 91.3%

Analysis of Variance

Source           DF      SS      MS       F      P
Regression        1  7.4068  7.4068  158.78  0.000
Residual Error   14  0.6531  0.0466
Total            15  8.0599
```

do el mismo método que se muestra en la sección 10.2. La única diferencia estriba en que se ajustará la línea de tendencia a los datos desestacionalizados en lugar de a los datos originales. Recuerde que para una tendencia lineal la ecuación de regresión estimada puede expresarse como

$$T_t = b_0 = b_1 t$$

donde

T_t = pronóstico de tendencia lineal en el periodo t

b_0 = intersección de la recta de tendencia lineal

b_1 = pendiente de la línea de tendencia

t = periodo

En la sección 10.2 se obtuvieron las fórmulas para calcular los valores de b_0 y b_1. Para ajustar una recta de tendencia lineal a los datos desestacionalizados de la tabla 10.21, el único cambio estriba en que al calcular b_0 y b_1 se utilizan los valores de las series de tiempo desestacionalizadas en lugar de los valores observados Y_t.

La figura 10.21 muestra los resultados por computadora que se obtuvieron con el procedimiento de análisis de regresión de *MINITAB* para estimar la línea de tendencia de las series de tiempo desestacionalizadas de ventas de televisores. La ecuación de tendencia lineal estimada es

$$\text{Ventas desestacionalizadas} = 5.10 = 0.148t$$

La pendiente de 0.148 indica que en los últimos 16 trimestres la empresa promedió un crecimiento desestacionalizado de las ventas de 148 televisores por trimestre. Si se supone que los datos de ventas de los últimos 16 trimestres es un indicador bastante bueno del futuro, esta ecuación se puede utilizar para proyectar la variación de tendencia de los próximos trimestres. Por ejemplo, si en esta ecuación se sustituye $t = 17$, obtenemos la proyección de la tendencia desestacionalizada para el siguiente trimestre, T_{17}.

$$T_{17} = 5.10 + 0.148(17) = 7.616$$

Por lo tanto, al utilizar los datos desestacionalizados, el pronóstico de tendencia lineal es 7.616 televisores para el próximo trimestre (periodo 17). Del mismo modo, los pronósticos de tendencia desestacionalizada para los próximos tres trimestres (periodos 18, 19 y 20) son 7 764, 7 912 y 8 060 televisores, respectivamente.

Ajustes estacionales

El último paso para formular un pronóstico cuando existe tanto una variación de tendencia como una variación estacional, es usar los índices estacionales a efecto de ajustar las proyecciones desestacionalizadas de tendencia. De vuelta al ejemplo de las ventas de televisores, tenemos una proyección desestacionalizada de la tendencia para los próximos cuatro trimestres.

TABLA 10.22 **Pronósticos trimestrales de la serie de tiempo de ventas de televisores**

Año	Trimestre	Pronóstico para la tendencia desestacionalizada	Índice estacional	Pronóstico trimestral
5	1	7616	0.93	(7616)(0.93) = 7083
	2	7764	0.84	(7764)(0.84) = 6522
	3	7912	1.09	(7912)(1.09) = 8624
	4	8060	1.14	(8060)(1.14) = 9188

Ahora es necesario ajustar el pronóstico para el efecto estacional. El índice estacional para el primer trimestre del año 5 ($t = 17$) es 0.93, por lo que se obtiene el pronóstico trimestral al multiplicar el pronóstico desestacionalizado basado en la tendencia ($T_{17} = 7\,616$) por el índice estacional (0.93). Por lo tanto, el pronóstico para el trimestre siguiente es 7 617(0.93) = 7 083. En la tabla 10.22 se presentan los pronósticos para los trimestres 17 a 20. El cuarto trimestre, con un alto volumen de ventas, tiene un pronóstico de 9 188 unidades, y el segundo trimestre, con un volumen bajo de ventas, tiene como pronóstico 6 522 unidades.

10.8 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

Elaboración de pronósticos con *MINITAB*

WEB file Gasoline

En esta sección se muestra cómo usar *MINITAB* para formular pronósticos mediante los siguientes métodos: promedios móviles, suavizamiento exponencial, proyección de tendencias, suavizamiento exponencial lineal de Holt y descomposición de una serie de tiempo.

Promedios móviles

Para mostrar cómo usar *MINITAB* en la elaboración de pronósticos con el método de promedios móviles, se calculará un pronóstico para las series de tiempo de venta de gasolina de la tabla 10.1 y de la figura 10.1. Los datos de ventas en las 12 semanas se ingresan en la columna 2 de la hoja de cálculo. Los siguientes pasos se utilizan para obtener un pronóstico del promedio móvil de tres semanas para la semana 13.

Paso 1. Seleccione el menú **Stat**.
Paso 2. Elija **Time Series**.
Paso 3. Elija **Moving Average**.
Paso 4. Cuando aparezca el cuadro de diálogo Moving Average:
Introduzca C2 en el cuadro **Variable**.
Introduzca 3 en el cuadro **MA length**.
Seleccione **Generate forecasts**.
Introduzca 1 en el cuadro **Number of forecasts**.
Introduzca 12 en el cuadro **Starting from origin**. Haga clic en **OK**.

Las medidas de exactitud del pronóstico para la semana 13 se muestran en la ventana de la sección. En los resultados de *MINITAB*, el error absoluto medio se etiqueta MAD y al cuadrado medio debido al error o error cuadrático medio, MSD.

Suavizamiento exponencial

WEB file Gasoline

Para mostrar cómo utilizar *MINITAB* con objeto de obtener un pronóstico de suavizamiento exponencial, se recurrirá nuevamente a los datos que se presentan en la tabla 10.1 y en la figura 10.1 a efecto de obtener un pronóstico de las ventas para la semana 13 de la serie de tiempo de las ventas de gasolina. Los datos de ventas para las 12 semanas se introducen en la columna 2 de la hoja de cálculo. Los siguientes pasos se utilizan para obtener un pronóstico para la semana 13 mediante una constante de suavizamiento de $\alpha = 0.2$.

Paso 1. Seleccione el menú **Stat**.
Paso 2. Elija **Time Series**.
Paso 3. Elija **Single Exp Smoothing**.
Paso 4. Cuando aparezca el cuadro de diálogo Single Exponential Smoothing:
Introduzca C2 en el cuadro **Variable**.
Seleccione la opción **Use** para Weight to Use in Smoothing.
Introduzca 0.2 en el cuadro Use.
Seleccione **Generate forecasts**.
Introduzca 1 en el cuadro **Number of forecasts**.
Introduzca 12 en el cuadro **Starting from origin**.
Seleccione **Options**.
Paso 5. Cuando aparezca el cuadro de diálogo Single Exponential Smoothing-Options:
Introduzca 1 en el cuadro **Use average of first K observations**.
Haga clic en **OK**.
Paso 6. Cuando aparezca el cuadro de diálogo Single Exponential Smoothing:
Haga clic en **OK**.

Las medidas de la exactitud del pronóstico y el pronóstico de suavizamiento exponencial de la semana 13 se muestran en la ventana de la sesión. En los resultados de *MINITAB*[†], el error absoluto medio se etiqueta MAD y el cuadrado medio debido al error o error cuadrático medio, MSD.

Proyección de tendencia

WEB file Bicycle

Para mostrar cómo permite *MINITAB* obtener pronósticos mediante la proyección de tendencias, se emplea el pronóstico para la serie de tiempo de ventas de bicicletas de la tabla 10.3 y de la figura 10.3. El número de años se introduce en la columna 1 y los datos de ventas en la columna 2 de la hoja de cálculo. Con los pasos siguientes se obtiene un pronóstico para el año 11 con la proyección de tendencia.

Paso 1. Seleccione el menú **Stat**.
Paso 2. Elija **Time Series**.
Paso 3. Elija **Trend Analysis**.
Paso 4. Cuando aparezca el cuadro de diálogo Trend Analysis:
Introduzca C2 en el cuadro **Variable**.
Elija **Linear** como tipo de modelo.
Seleccione **Generate forecasts**.
Introduzca 1 en el cuadro **Number of forecasts**.
Introduzca 10 en el cuadro **Starting from origin**.
Haga clic en **OK**.

La ecuación para la tendencia lineal de las medidas de exactitud del pronóstico y los pronósticos para el año siguiente se muestran en la ventana de la sesión. En los resultados de *MINITAB*, el error absoluto medio se etiqueta MAD y el error cuadrático medio MSD. Para obtener pronósticos de una tendencia cuadrática o una tendencia exponencial, seleccione **Quadratic** en **Exponential growth** en vez de **Linear** en el paso 4.

Suavizamiento exponencial lineal de Holt

WEB file Bicycle

Para mostrar cómo utilizar *MINITAB* para elaborar pronósticos con el método de suavizamiento exponencial lineal de Holt, se desarrolla nuevamente un pronóstico para la serie de tiempo de ventas de bicicletas de la tabla 10.3 y la figura 10.3. En *MINITAB*, el método de suavizamiento

[†] El valor de MSD que proporciona *MINITAB* no es el mismo que el valor del CME. *MINITAB* utiliza 17 como pronóstico para la semana 1, así que para calcular el MSD utiliza los datos de las 12 semanas. En la sección 10.3 se calcula el CME utilizando sólo los datos para las semanas 2 a 12 porque no se contaba con valores pasados que permitieran obtener un pronóstico para la semana 1.

exponencial lineal de Holt se refiere como suavizamiento exponencial doble (Double Exponential Smoothing). El número de años se introduce en la columna 1 y los datos de ventas en la columna 2 de la hoja de cálculo. Los pasos siguientes pueden utilizarse para pronosticar las ventas del año 11 mediante el suavizamiento exponencial lineal de Holt con $\alpha = 0.1$ y $\beta = 0.2$.

Paso 1. Seleccione el menú **Stat**.
Paso 2. Elija **Time Series**.
Paso 3. Elija **Double Exp Smoothing**.
Paso 4. Cuando aparezca el cuadro de diálogo Double Exponential Smoothing:
Introduzca C2 en el cuadro **Variable**.
Seleccione la opción **Use** para Weights to Use in Smoothing.
Introduzca 0.1 en el cuadro **level**.
Introduzca 0.2 en el cuadro **trend**.
Seleccione **Generate forecasts**.
Introduzca 1 en el cuadro **Number of forecasts**.
Introduzca 10 en el cuadro **Starting from origin**.
Haga clic en **Ok**.

Las medidas de exactitud de los pronósticos y los pronósticos de suavizamiento exponencial lineal de Holt para el año 11 se muestran en la ventana de la sesión. En los resultados de *MINITAB*, el error absoluto medio se etiqueta MAD y el cuadrado medio debido al error o error cuadrático medio, MSD.

Descomposición de una serie de tiempo

WEB file TVSales

El uso de *MINITAB* para pronosticar una serie de tiempo con tendencia y estacionalidad mediante la descomposición de series de tiempo se muestra elaborando un pronóstico para la serie de tiempo de las ventas de televisores de la tabla 10.6 y de la figura 10.6. En *MINITAB*, el usuario tiene la opción de elegir un modelo de descomposición multiplicativa o aditiva. En la sección 10.7 se ilustra cómo utilizar el método multiplicativo. El número de años se introduce en la columna 1, los valores trimestrales en la columna 2 y los datos de las ventas en la columna 3 de la hoja de cálculo. Con los siguientes pasos se obtiene un pronóstico para el próximo trimestre.

Paso 1. Seleccione el menú **Stat**.
Paso 2. Elija **Time Series**.
Paso 3. Elija **Decomposition**.
Paso 4. Cuando aparezca el cuadro de diálogo Decomposition:
Introduzca C3 en el cuadro **Variable**.
Introduzca 4 en el cuadro **Season Length**.
Seleccione **Multiplicative** para Method Type.
Elija **Trend plus Seasonal** para Model Components.
Seleccione **Generate forecasts**.
Introduzca 1 en el cuadro **Number of forecasts**.
Introduzca 16 en el cuadro **Starting from origin**. Haga clic en **OK**.

Los índices estacionales,† las medidas de exactitud de los pronósticos y los pronósticos para el próximo trimestre se muestran en la ventana de la sesión. En los resultados de *MINITAB*, el error absoluto medio se etiqueta MAD y el cuadrado medio debido al error o error cuadrático medio, MSD.

†Los resultados difieren ligeramente de los que se muestran en la tabla 10.7 debido a que *MINITAB* calcula los índices estacionales con la mediana de los valores estacional-irregular.

EJERCICIOS

TÉCNICAS BÁSICAS

10.1 Considere los siguientes datos de una serie de tiempo.

t	1	2	3	4	5
Y_t	6	11	9	14	15

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Desarrolle la ecuación de tendencia lineal para esta serie de tiempo.

c. ¿Cuál es el pronóstico para $t = 6$?

10.2 Considere la siguiente serie de tiempo.

t	1	2	3	4	5	6	7
Y_t	82	60	44	35	30	29	35

a. Elabore una gráfica de serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Utilice *MINITAB* o Excel para desarrollar la ecuación de tendencia cuadrática para la serie de tiempo.

c. ¿Cuál es el pronóstico para $t = 8$?

10.3 El presidente de una pequeña empresa de manufactura está preocupado por el continuo incremento en los costos de fabricación de los últimos años. Las siguientes cifras presentan una serie de tiempo del costo unitario del producto principal de la empresa en los últimos ocho años.

Año	Costo unitario ($)	Año	Costo unitario ($)
1	20.00	5	26.60
2	24.50	6	30.00
3	28.20	7	31.00
4	27.50	8	36.00

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Desarrolle la ecuación de tendencia lineal para esta serie de tiempo.

c. ¿Qué porcentaje han aumentado cada año los costos promedio de la empresa?

d. Proporcione un cálculo estimado del costo unitario para el próximo año.

10.4 Considere la siguiente serie de tiempo.

Trimestre	Año 1	Año 2	Año 3
1	71	68	62
2	49	41	51
3	58	60	53
4	78	81	72

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Utilice las siguientes variables ficticias para desarrollar una ecuación de regresión estimada que tome en cuenta los efectos estacionales en los datos: Qtr1 = 1 si es el trimestre 1, 0 en caso contrario; Qtr2 = 1 si es el trimestre 2, 0 en caso contrario; Qtr 3 = 1 si es el trimestre 3, 0 en caso contrario.

c. Formule los pronósticos trimestrales para el año siguiente.

10.5 Considere los datos de la siguiente serie de tiempo.

Semana	1	2	3	4	5	6
Valor	18	13	16	11	17	14

Utilizando el método ingenuo (el valor más reciente) como pronóstico para la semana próxima, calcule las siguientes medidas de exactitud de los pronósticos.

a. Error absoluto medio.

b. Cuadrado medio debido al error.

c. Error porcentual absoluto medio.

d. ¿Cuál es el pronóstico para la semana 7?

10.6 Considere los siguientes datos de la serie de tiempo.

Mes	1	2	3	4	5	6	7
Valor	24	13	20	12	19	23	15

a. Calcule el valor del CME utilizando el valor más reciente como pronóstico para el periodo próximo. ¿Cuál es el pronóstico para el mes 8?

b. Calcule el valor del CME utilizando el promedio de todos los datos disponibles como pronóstico para el siguiente periodo. ¿Cuál es el pronóstico para el mes 8?

c. ¿Qué método parece proveer el mejor pronóstico?

10.7 Considere los datos de la siguiente serie de tiempo.

Trimestre	Año 1	Año 2	Año 3
1	4	6	7
2	2	3	6
3	3	5	6
4	5	7	8

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Proporcione los valores del promedio móvil de cuatro trimestres y los valores del promedio móvil centrado para esta serie de tiempo.

c. Calcule los índices estacionales y los índices estacionales ajustados para los cuatro trimestres.

APLICACIONES

10.8 Textos universitarios I Los datos de las ventas trimestrales (número de ejemplares vendidos) de un libro de texto universitario en los últimos tres años son los siguientes:

Quarter	Year 1	Year 2	Year 3
1	1 690	1 800	1 850
2	940	900	1 100
3	2 625	2 900	2 930
4	2 500	2 360	2 615

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Utilice las siguientes variables ficticias para desarrollar una ecuación de regresión estimada que tome en cuenta los efectos estacionales de los datos: Qrt = 1 si el trimestre es 1, 0 en caso contrario; Qtr2 = 1 si el trimestre es 2, 0 en caso contrario; Qtr3 = 1 si el trimestre es 3, 0 en caso contrario.

c. Elabore los pronósticos trimestrales para el próximo año.

d. Suponga que $t = 1$ se refiere a la observación del trimestre 1 del año 1; $t = 2$ a la observación del trimestre 2 del año 1; . . . y $t = 12$ a la observación del trimestre 4 del año 3.

Utilice las variables ficticias definidas en la parte **b**. así como t, para desarrollar una ecuación de regresión estimada que tome en cuenta los efectos estacionales y cualquier tendencia lineal en la serie de tiempo. A partir de los efectos estacionales en los datos y la tendencia lineal, calcule los pronósticos trimestrales para el próximo año.

10.9 Mantenimiento de edificios Los gastos del mantenimiento del césped ($) por mes (Month) para un edificio de seis departamentos en el sur de la Florida durante tres años (Year) son los siguientes:

Month	Year 1	Year 2	Year 3
January	170	180	195
February	180	205	210
March	205	215	230
April	230	245	280
May	240	265	290
June	315	330	390
July	360	400	420
August	290	335	330
September	240	260	290
October	240	270	295
November	230	255	280
December	195	220	250

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Desarrolle una ecuación de regresión estimada que tome en cuenta cualquier efecto de estacionalidad o de tendencia lineal en los datos. Utilice las siguientes variables ficticias para tomar en cuenta los efectos estacionales: Jan = 1 si es enero, 0 en caso contrario; Feb = 1 si es febrero, 0 en caso contrario; Mar = 1 si es marzo, 0 en caso contrario; . . . Nov = 1 si es noviembre, 0 en caso contrario. Observe que al utilizar este método de codificación, cuando las 11 variables ficticias son 0, la observación corresponde a un gasto en diciembre.

c. Calcule los pronósticos mensuales para el siguiente año con base en los efectos tanto de tendencia como estacionales.

10.10 Textos universitarios II A continuación se presentan los datos de las ventas por trimestre (Quarter) del número de ejemplares vendidos para un libro de texto universitario en los últimos tres años (Year 1, 2 y 3).

Quarter	Year 1	Year 2	Year 3	
1	1 690	1	800	1 850
2	940	900	1	100
3	2 625	2	900	2 930
4	2 500	2	360	2 615

a. Elabore una gráfica de la serie de tiempo. ¿Qué tipo de patrón existe en los datos?

b. Para esta serie de tiempo, proporcione los valores del promedio móvil de cuatro trimestres y del promedio móvil centrado.

c. Calcule los índices estacionales y los índices estacionales ajustados de los cuatro trimestres.

d. ¿Cuándo obtiene la editorial el mayor índice estacional? ¿Parece razonable este resultado? Explique.

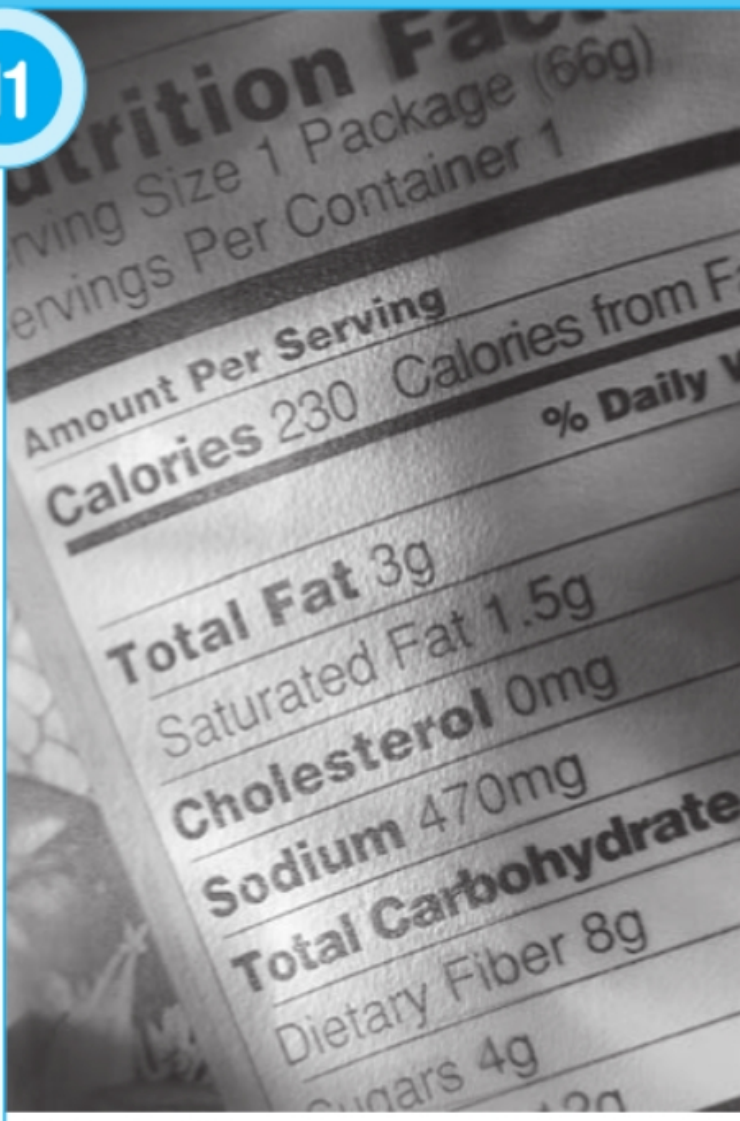
e. Desestacionalice la serie de tiempo.

f. Calcule la ecuación de tendencia lineal para los datos desestacionalizados y pronostique las ventas utilizando la ecuación de tendencia lineal.

g. Modifique los pronósticos de tendencia lineal utilizando los índices estacionales ajustados calculados en la parte **c**.

11

Estadística no paramétrica



© Don Carstens/Brand X/CORBIS

OBJETIVOS GENERALES

En los capítulos 3 a 5 presentamos técnicas estadísticas para comparar dos poblaciones contrastando sus respectivos parámetros poblacionales (por lo general, sus medias poblacionales). Las técnicas de los capítulos 3 y 5 son aplicables a los datos que son al menos cuantitativos y las del capítulo 4 son aplicables a los datos que tienen distribuciones normales. El propósito de este capítulo es presentar varias pruebas estadísticas útiles en la comparación de poblaciones para los numerosos tipos de datos que no satisfagan las suposiciones especificadas en los capítulos 3 a 5.

ÍNDICE DEL CAPÍTULO

- El coeficiente de correlación de rango
- La prueba del rango con signo de Wilcoxon para un experimento pareado
- La prueba del signo para un experimento pareado
- La prueba de la suma de rango de Wilcoxon: muestras aleatorias independientes
- La prueba F_r de Friedman
- Pruebas paramétricas contra no paramétricas

11.1 INTRODUCCIÓN, PRUEBAS NO PARAMÉTRICAS

Algunos experimentos generan respuestas que pueden ser ordenadas o clasificadas, pero no es posible medir numéricamente el valor real de la respuesta excepto con una escala arbitraria que usted cree. Es probable que usted sólo pueda decir si una observación es mayor que otra. Quizá sea capaz de clasificar todo un conjunto de observaciones sin saber en realidad los valores numéricos exactos de las mediciones. Revisemos algunos ejemplos:

- Las habilidades de ventas de cuatro vendedores son clasificadas de la mejor a la peor.
- Las características de comestible y sabor de cinco marcas de fibra de pasitas se clasifican en una escala arbitraria de 1 a 5.
- Cinco diseños de automóviles se clasifican del más atractivo al menos atractivo.

¿Cómo se analizan estos tipos de datos? Los métodos estadísticos de muestra pequeña presentados en capítulos anteriores son válidos sólo cuando la(s) población(es) muestreada(s) es (son) normal(es) o aproximadamente lo es (son). Los datos formados por rangos o escalas arbitrarias de 1 a 5 *no satisfacen la suposición de normalidad*, incluso a un grado razonable. En algunas aplicaciones, las técnicas son válidas si las muestras se toman al azar de poblaciones cuyas varianzas son iguales.

Cuando los datos no parecen satisfacer estas y suposiciones similares, es posible usar un método alternativo, **métodos estadísticos no paramétricos**. Los métodos no paramétricos por lo general especifican las hipótesis en términos de distribuciones poblacionales más que parámetros como medias y desviaciones estándar. Es frecuente que las suposiciones paramétricas sean sustituidas por suposiciones más generales acerca de las distribuciones poblacionales y las clasificaciones de las observaciones se usen a veces en lugar de las mediciones reales.

Investigaciones realizadas demuestran que las pruebas estadísticas no paramétricas son tan capaces de detectar diferencias entre poblaciones como los métodos paramétricos de capítulos anteriores, cuando se satisfacen la normalidad y otras suposiciones. Pueden ser, y con frecuencia son, más potentes para detectar diferencias poblacionales cuando estas suposiciones no se satisfacen. Por esta razón, algunos estadísticos están a favor de usar procedimientos no paramétricos en lugar de sus similares paramétricos.

Presentaremos métodos no paramétricos apropiados para comparar dos o más poblaciones usando ya sea muestras independientes o pareadas. También se presentará una medida de asociación que es útil para determinar si una variable aumenta cuando la otra aumenta o si una variable disminuye cuando la otra aumenta.

11.2 PRUEBA DE LOS SIGNOS

La prueba del signo es un procedimiento bastante sencillo que se usa para comparar dos poblaciones cuando las muestras consisten en observaciones pareadas. Este tipo de diseño experimental se denomina diseño de **diferencia pareada** o **pares comparados**, que en la sección 4.4 se usó para comparar el promedio de desgaste para dos tipos de llantas. En general, para cada par se mide si la primera respuesta, por ejemplo A, excede a la segunda respuesta, por ejemplo a B. El estadístico de prueba es x, el número de veces que A excede a B en los **n** pares de observaciones.

Cuando las dos distribuciones poblacionales son idénticas, la probabilidad de que A exceda a B es igual a $p = .5$, y x, el número de veces que A excede a B, tiene una distribución binomial. Sólo pares sin empates se incluyen en la prueba. En consecuencia, se puede probar la hipótesis de distribuciones poblacionales idénticas al probar H_0: $p = .5$ contra una alternativa ya sea de una o dos colas. Los valores críticos para la región de rechazo o valores p exactos se hallan usando las tablas binomiales acumulativas del Apéndice.

LA PRUEBA DEL SIGNO PARA COMPARAR DOS POBLACIONES

1. Hipótesis nula: H_0: Las dos distribuciones poblacionales son idénticas y P (A excede a B) $= p = .5$
2. Hipótesis alternativa:
   a. H_a: Las distribuciones poblacionales no son idénticas y $p \neq .5$
   b. H_a: La población de A mediciones se corre a la derecha de la población de B mediciones y $p > .5$
   c. H_a: La población de A mediciones se corre a la izquierda de la población de B mediciones y $p < .5$
3. Estadístico de prueba: Para n, el número de pares sin empates, use x, el número de veces que (A – B) es positivo.
4. Región de rechazo:
   a. Para la prueba de dos colas H_a: $p \neq .5$, rechazar H_0 si $x \leq x_L$ o $x \geq x_U$, donde $P(x \leq x_L) \leq \alpha/2$ y $P(x \geq x_U) \leq \alpha/2$ para x que tenga una distribución binomial con $p = .5$.
   b. Para H_a: $p > .5$, rechazar H_0 si $x \geq x_U$ con $P(x \geq x_U) \leq \alpha$.
   c. Para H_a: $p < .5$, rechazar H_0 si $x \leq x_L$ con $P(x \leq x_L) < \alpha$.

O calcule el valor p y rechace H_0 si el valor $p < \alpha$.

Un problema que probablemente ocurra cuando se realiza una prueba del signo es que las mediciones asociadas con uno o más pares pueden ser iguales y, por tanto, resultan en **observaciones empatadas**. Cuando esto ocurra, elimine los pares empatados y reduzca n, el número total de pares. El ejemplo siguiente le ayudará a entender cómo se construye y utiliza la prueba del signo.

EJEMPLO 11.1

Los números de fusibles eléctricos defectuosos producidos por dos líneas de producción, A y B, se registraron a diario durante un periodo de 10 días, los resultados aparecen en la tabla 11.1. La variable de respuesta, el número de fusibles defectuosos, tiene una distribución binomial exacta con un gran número de fusibles producidos por día. Aun cuando esta variable tendrá aproximadamente una distribución normal, el supervisor de planta preferiría una prueba estadística rápida y fácil para determinar si una línea de producción tiende a producir más fusibles defectuosos que la otra. Use la prueba del signo para probar la hipótesis apropiada.

TABLA 11.1 Fusibles defectuosos de dos líneas de producción

Día	Línea A	Línea B	Signo de diferencia
1	170	201	–
2	164	179	–
3	140	159	–
4	184	195	–
5	174	177	–
6	142	170	–
7	191	183	+
8	169	179	–
9	161	170	–
10	200	212	–

Solución Para este experimento de *diferencia pareada*, x es el número de veces que la observación de la línea A excede la de la línea B en un día determinado. Si no hay diferencia en las distribuciones de fusibles defectuosos para las dos líneas de producción, entonces p, la proporción de días en los que A excede a B, es .5, que es el valor hipotético en una prueba del

parámetro binomial p. Valores muy pequeños o muy grandes de x, el número de veces que A excede de B, son contrarios a la hipótesis nula.

Como $n = 10$ y el valor hipotético de p es .5, la tabla 1 del Apéndice se puede usar para hallar el valor p exacto para la prueba de

H_0: $p = .5$ contra H_a: $p \neq .5$

El valor observado del estadístico de prueba, que es el número de signos "más" en la tabla, es $x = 1$, y el valor p se calcula como

$$\text{valor } p = 2P(x \leq 1) = 2(.011) = .022$$

El valor p bastante pequeño = .022 permite rechazar H_0 al nivel de 5%. Hay evidencia significativa para indicar que el número de fusibles defectuosos no es el mismo para las dos líneas de producción; de hecho, la línea B produce más fusibles defectuosos que la A. En este ejemplo, la prueba del signo es una herramienta aproximada, fácil de calcular, para detectar líneas de producción con falla y funciona perfectamente bien para detectar una diferencia significativa usando sólo una cantidad mínima de información.

Aproximación normal para la prueba del signo

Cuando el número de pares n es grande, los valores críticos para el rechazo de H_0 y los valores p aproximados se pueden hallar usando una aproximación normal a la distribución de x. Debido a que la distribución binomial es perfectamente simétrica cuando $p = .5$, esta aproximación funciona muy bien, incluso para n de sólo 10.

Para $n \geq 25$, se puede efectuar la prueba del signo usando el estadístico z,

$$z = \frac{x - np}{\sqrt{npq}} = \frac{x - .5n}{.5\sqrt{n}}$$

como el estadístico de prueba. Al usar z, se prueba la hipótesis nula $p = .5$ contra la alternativa $p \neq .5$ para una prueba de dos colas o contra la alternativa $p > .5$ (o $p < .5$) para una prueba de una cola.

PRUEBA DEL SIGNO PARA MUESTRAS GRANDES: $n \geq 25$

1. La hipótesis nula: H_0: $p = .5$ (un tratamiento no se prefiere a un segundo tratamiento).
2. Hipótesis alternativa: H_a: $p \neq .5$, para una prueba de dos colas (NOTA: Usamos la prueba de dos colas como ejemplo. Muchos análisis podrían requerir una prueba de una cola.)
3. Estadístico de prueba: $z = \dfrac{x - .5n}{.5\sqrt{n}}$
4. Región de rechazo: Rechace H_0 si $z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$, donde $z_{\alpha/2}$ es el valor z de la tabla 3 del Apéndice correspondiente a un área de $\alpha/2$ en la cola superior de la distribución normal.

EJEMPLO 11.2 Un superintendente de producción dice que no hay diferencia entre los porcentajes de accidentes de empleados en turnos de día o de noche en una gran planta manufacturera. Se registra el número diario de accidentes para los turnos de día y de noche durante $n = 100$ días. Se encuentra que el número diario de accidentes en el turno de noche x_E excedió al número correspondiente de éstos en el turno de día x_D en 63 de los 100 días. ¿Estos resultados dan suficiente evidencia para indicar que más accidentes tienden a ocurrir en un turno que en el otro, o bien, es equivalente, que $P(x_E > x_D) \neq 1/2$?

Solución Este estudio es un experimento de diferencia pareada, con $n = 100$ pares de observaciones correspondientes a los 100 días. Para probar la hipótesis nula de que las dos distribuciones de accidentes son idénticas, se puede usar el estadístico de prueba

$$z = \frac{x - .5n}{.5\sqrt{n}}$$

donde x es la cantidad de días en los cuales el número de accidentes en el turno de noche excedió al de accidentes en el turno de día. Entonces, para $\alpha = .05$, se puede rechazar la hipótesis nula si $z \geq 1.96$ o $z \leq -1.96$. Sustituyendo en la fórmula para z, se obtiene

$$z = \frac{x - .5n}{.5\sqrt{n}} = \frac{63 - (.5)(100)}{.5\sqrt{100}} = \frac{13}{5} = 2.60$$

Como el valor calculado de z excede de $z_{\alpha/2} = 1.96$, se puede rechazar la hipótesis nula. Los datos dan suficiente evidencia para indicar una diferencia en las distribuciones del porcentaje de accidentes para el turno de día contra el de noche.

¿Cuándo debe usarse la prueba del signo en vez de la prueba t? Cuando se da sólo la *dirección* de la diferencia en la medición, *sólo* se puede usar la prueba del signo. Por el contrario, cuando los datos son cuantitativos y satisfacen las suposiciones de normalidad y varianza constante, debe usarse la prueba t pareada. Se utiliza una gráfica de probabilidad normal para evaluar normalidad, en tanto que una gráfica de los residuales $(d_i - \bar{d})$ puede revelar grandes desviaciones que indicarían una varianza diferente de un par a otro. Cuando haya dudas acerca de la validez de las suposiciones, los estadísticos recomiendan con frecuencia que se realicen ambas pruebas. Si las dos llegan a las mismas conclusiones, entonces los resultados de la prueba paramétrica se consideran válidos.

11.3 PRUEBA DE LA SUMA DE RANGOS

Al comparar las medias de dos poblaciones basadas en muestras independientes, el estadístico pivote era la diferencia en las medias muestrales. Si no está seguro de que las suposiciones requeridas para una prueba t de dos muestras sea satisfecha, una alternativa es sustituir los valores de las observaciones por sus rangos y proceder como si éstos fueran las observaciones reales. Dos pruebas no paramétricas diferentes usan un estadístico de prueba basado en estos rangos de muestra:

- Prueba de suma de rango de Wilcoxon
- Prueba U de Mann-Whitney

Son *equivalentes* en que usan la misma información muestral. El procedimiento que presentaremos es la prueba de suma de rango de Wilcoxon, propuesta por Frank Wilcoxon, que está basada en la suma de los rangos de la muestra que tiene el tamaño muestral más pequeño.

Suponga que tenemos n_1 observaciones de la población 1 y n_2 observaciones de la población 2. La hipótesis nula a probar es que las dos distribuciones poblacionales son idénticas, contra la hipótesis alternativa de que las distribuciones poblacionales son diferentes. Éstas son las posibilidades para las dos poblaciones:

- Si H_0 es verdadera y las observaciones han provenido de las mismas o idénticas poblaciones, entonces las observaciones de ambas muestras deben mezclarse al azar cuando conjuntamente sean de rango de pequeño a grande. La suma de los rangos de las observaciones de la muestra 1 debe ser similar a la suma de los rangos de la muestra 2.

- Si, por el contrario, las observaciones de la población 1 tienden a ser más pequeñas que las de la población 2, entonces estas observaciones tendrían los rangos más pequeños porque la mayoría de éstas serían más pequeñas que las de la población 2. La suma de los rangos de ellas sería "pequeña".
- Si las observaciones de la población 1 tienden a ser más grandes que las de la población 2, se les asignarían rangos más grandes. La suma de los rangos de estas últimas tendería a ser "grande".

Por ejemplo, suponga que tenemos $n_1 = 3$ observaciones de la población 1, es decir, 2, 4 y 6, y $n_2 = 4$ observaciones de la población 2, o sea 3, 5, 8 y 9. La tabla 11.2 muestra siete observaciones ordenadas de pequeñas a grandes.

TABLA 11.2 **Siete observaciones en orden**

Observación	x_1	y_1	x_2	y_2	x_3	y_3	y_4
Datos	2	3	4	5	6	8	9
Rango	1	2	3	4	5	6	7

A la observación más pequeña, $x_1 = 2$, se le asigna el rango 1; a la siguiente observación más pequeña, $y_1 = 3$, se le asigna el rango 2, y así sucesivamente. La *suma de los rangos* de las observaciones de la muestra 1 es $1 + 3 + 5 = 9$ y la **suma de rangos** de la muestra 2 es $2 + 4 + 6 + 7 = 19$. ¿Cómo se determina si la suma de rangos de las observaciones de la muestra 1 es significativamente pequeña o significativamente grande? Esto depende de la distribución de probabilidad de la suma de rangos de una de las muestras. Como los rangos para $n_1 + n_2 = N$ observaciones son los primeros N enteros, se puede demostrar que la suma de estos rangos es $N(N + 1)/2$. En este sencillo ejemplo, la suma de $N = 7$ rangos es $1 + 2 + 3 + 4 + 5 + 6 + 7 = 7(8)/2$ o sea 28. En consecuencia, si el experimentador conoce la suma de rangos para una de las muestras, puede hallar la otra por sustracción. En nuestro ejemplo, observe que la suma de rangos para la muestra 1 es 9, en tanto que la segunda suma de rangos es $(28 - 9) = 19$. Esto significa que sólo una de las dos sumas de rangos es necesaria para la prueba. Para simplificar la tabulación de valores críticos para esta prueba debe usarse la suma de rangos de la muestra más pequeña como estadístico de prueba. ¿Qué ocurre si dos o más observaciones son iguales? A observaciones empatadas se les asigna el promedio de los rangos que tendrían las observaciones si hubieran sido ligeramente diferentes en valor.

Para poner en práctica la prueba de la suma de rangos de Wilcoxon, supongamos que muestras aleatorias independientes de tamaño n_1 y n_2 se seleccionan de las poblaciones 1 y 2, respectivamente. Representemos con n_1 al menor de los dos tamaños muestrales y con T_1 la suma de los rangos de las observaciones de la muestra 1. Si la población 1 está a la izquierda de la población 2, T_1 será "pequeña". T_1 será "grande" si la población 1 está a la derecha de la 2.

FÓRMULAS PARA EL ESTADÍSTICO DE LA SUMA DE RANGOS DE WILCOXON (PARA MUESTRAS INDEPENDIENTES)

Sean

T_1 = Suma de rangos para la primera muestra

$T_1^* = n_1(n_1 + n_2 + 1) - T_1$

T_1^* es el valor de la suma de rangos para n_1 si las observaciones se hubieran ordenado de *grande a pequeña*. (*No* es la suma de rangos para la segunda muestra.) Dependiendo de la naturaleza de la hipótesis alternativa, uno de estos dos valores se elegirá como estadístico de prueba, T.

La tabla 7 del Apéndice se puede usar con el fin de localizar *valores críticos* para el estadístico de prueba para cuatro valores diferentes de pruebas de una cola con $\alpha = .05, .025, .01$ y $.005$. Para usar la tabla 7 del Apéndice para una prueba de dos colas, los valores de α se duplican, es decir, $\alpha = .10, .05, .02$ y $.01$. La entrada de la tabla da el valor de a tal que $P(T \leq a) \leq \alpha$. Con la intención de ver cómo localizar un valor crítico para la prueba de suma de rango de Wilcoxon, suponga que $n_1 = 8$ y $n_2 = 10$ para una prueba de una cola con $\alpha = .05$. Se puede usar la tabla 7a del Apéndice, una parte de la cual se reproduce en la tabla 11.3. Observe que la tabla está construida suponiendo que $n_1 \leq n_2$. Es por esta razón que designamos la población con el tamaño muestral más pequeño como población 1. Los valores de n_1 se muestran en sentido horizontal en la parte superior de la tabla y los de n_2 se muestran en sentido vertical al lado izquierdo. La entrada $a = 56$, sombreada, es el valor crítico para rechazar H_0. La hipótesis nula de igualdad de las dos distribuciones debe ser rechazada si el valor observado del estadístico de prueba T es menor o igual a 56.

TABLA 11.3 **Parte de los valores críticos de cola izquierda al 5%, tabla 7 del Apéndice**

	n_1						
n_2	2	3	4	5	6	7	8
3	—	6					
4	—	6	11				
5	3	7	12	19			
6	3	8	13	20	28		
7	3	8	14	21	29	39	
8	4	9	15	23	31	41	51
9	4	10	16	24	33	43	54
10	4	10	17	26	35	45	56

PRUEBA DE LA SUMA DE RANGO DE WILCOXON

Con n_1 denotemos la más pequeña de las dos muestras. Esta muestra proviene de la población 1. Las hipótesis a probar son

H_0: Las distribuciones para las poblaciones 1 y 2 son idénticas

contra una de tres hipótesis alternativas:

H_a: Las distribuciones para las poblaciones 1 y 2 son diferentes (una prueba de dos colas)
H_a: La distribución para la población 1 está a la izquierda de aquella de la población 2 (una prueba de cola izquierda)
H_a: La distribución para la población 1 está a la derecha de aquella de la población 2 (una prueba de cola derecha)

Procedimiento

1. Ordene todas las $n_1 + n_2$ observaciones de pequeña a grande.
2. Encuentre T_1, la suma de rangos para las observaciones de la muestra 1. Éste es el estadístico de prueba para la prueba de cola izquierda.
3. Encuentre $T_1^* = n_1(n_1 + n_2 + 1) - T_1$, la suma de los rangos de las observaciones de la población 1 si los rangos asignados se hubieran invertido de grandes a pequeños. (El valor de T_1^* no es la suma de los rangos de las observaciones de la muestra 2.) Éste es el estadístico de prueba para una prueba de cola derecha.
4. El estadístico de prueba para una prueba de dos colas es T, la *mínima* de T_1 y T_1^*.
5. H_0 es rechazada si el estadístico de prueba observado es menor o igual al valor crítico hallado usando la tabla 7 del Apéndice.

Ilustramos el uso de la tabla 7 con el siguiente ejemplo.

EJEMPLO 11.3 Las frecuencias de aleteo de dos especies de abejas Euglossine fueron registradas para una muestra de $n_1 = 4$ *Euglossa mandibularis* Friese (especie 1) y $n_2 = 6$ *Euglossa imperialis* Cockerell (especie 2). Las frecuencias se detallan en la tabla 11.4. ¿Puede usted concluir que las distribuciones de aleteo difieren para estas dos especies? Pruebe usando $\alpha = .05$.

TABLA 11.4 Frecuencias de aleteo para dos especies de abejas

Especie 1	Especie 2
235	180
225	169
190	180
188	185
	178
	182

Solución Primero es necesario ordenar las observaciones de pequeña a grande, como se muestra en la tabla 11.5.

TABLA 11.5 Frecuencias de aleteo ordenadas de pequeña a grande

Datos	Especie	Orden
169	2	1
178	2	2
180	2	3
180	2	4
182	2	5
185	2	6
188	1	7
190	1	8
225	1	9
235	1	10

Las hipótesis a probar son

H_0: Las distribuciones de las frecuencias de aleteo son las mismas para las dos especies

contra

H_a: Las distribuciones de las frecuencias de aleteo difieren para las dos especies

Como el tamaño muestral para individuos de la especie 1, $n_1 = 4$, es el más pequeño de los dos tamaños muestrales, tenemos

$$T_1 = 7 + 8 + 9 + 10 = 34$$

y

$$T_1^* = n_1(n_1 + n_2 + 1) - T_1 = 4(4 + 6 + 1) - 34 = 10$$

Para una prueba de dos colas, el estadístico de prueba es $T = 10$, la menor de $T_1 = 34$ y $T_1^* = 10$.

Para esta prueba de dos colas con $\alpha = .05$, puede usar la tabla 7b del Apéndice con $n_1 = 4$ y $n_2 = 6$. El valor crítico de T tal que $P(T \leq a) \leq \alpha/2 = .025$ es 12, y se debe rechazar la hipótesis nula si el valor observado de T es 12 o menos. Como el valor observado del estadístico de prueba, $T = 10$, es menor que 12, se puede rechazar la hipótesis de distribuciones de frecuencias de aleteo iguales al nivel de significancia de 5%.

Una salida impresa *MINITAB* de la prueba de suma de rango de Wilcoxon (llamada Mann-Whitney por *MINITAB*) para estos datos se da en la figura 11.1. Al final de este capítulo se encuentran instrucciones para generar esta salida impresa en la sección 11.6. Observe que la suma de rango de la primera muestra se da como W = 34.0, que concuerda con nuestros cálculos. Con un valor p reportado de .0142 calculado por *MINITAB*, se puede rechazar la hipótesis nula al nivel de 5%.

FIGURA 11.1
Salida impresa para el ejemplo 11.3

Prueba Mann-Whitney y CI: especie 1, especie 2

```
               N     Median
Species 1      4     207.50
Species 2      6     180.00

Point estimate for ETA1-ETA2 is 30.50
95.7 Percent CI for ETA1-ETA2 is (5.99,56.01)
W = 34.0
Test of ETA1 = ETA2 vs ETA1 not = ETA2 is significant at 0.0142
The test is significant at 0.0139 (adjusted for ties)
```

Aproximación normal para la prueba de suma de rango de Wilcoxon

La tabla 7 del Apéndice contiene valores críticos para tamaños muestrales de $n_1 \leq n_2 = 3, 4, \ldots, 15$. Siempre que n_1 no sea demasiado pequeña,[†] las aproximaciones a las probabilidades para el estadístico T de la suma de rango de Wilcoxon se pueden hallar usando una aproximación normal a la distribución de T. Se puede demostrar que la media y la varianza de T son

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2} \quad \text{y} \quad \sigma_T^2 = \frac{n_1 n_2(n_1 + n_2 + 1)}{12}$$

La distribución de

$$z = \frac{T - \mu_T}{\sigma_T}$$

es aproximadamente normal con media 0 y desviación estándar 1 para valores de n_1 y n_2 de sólo 10.

Si se intenta esta aproximación para el ejemplo 11.1, se obtiene

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{4(4 + 6 + 1)}{2} = 22$$

y

$$\sigma_T^2 = \frac{n_1 n_2(n_1 + n_2 + 1)}{12} = \frac{4(6)(4 + 6 + 1)}{12} = 22$$

El valor p para esta prueba es $2P(T \geq 34)$. Si se usa una corrección de .5 para continuidad al calcular el valor de z porque n_1 y n_2 son pequeñas,[‡] tenemos

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{(34 - .5) - 22}{\sqrt{22}} = 2.45$$

El valor p para esta prueba es

$$2P(T \geq 34) \approx 2P(z \geq 2.45) = 2(.0071) = .0142$$

el valor reportado en la salida impresa *MINITAB* de la figura 11.1.

[†] Algunos investigadores indican que la aproximación normal es adecuada para muestras de hasta sólo $n_1 = n_2 = 4$.

[‡] Como el valor de $T = 34$ está a la derecha de la media 22, la resta de .5 al usar la aproximación normal toma en cuenta el límite inferior de la barra arriba del valor 34 en la distribución de probabilidad de T.

PRUEBA DE LA SUMA DE RANGO DE WILCOXON PARA MUESTRAS GRANDES: $n_1 \geq 10$ Y $n_2 \geq 10$

1. Hipótesis nula: H_0: Las distribuciones de las poblaciones son idénticas.
2. Hipótesis alternativa: H_a: Las dos distribuciones poblacionales no son idénticas (una prueba de dos colas); o H_a: La distribución de la población 1 se corre a la derecha (o a la izquierda) de la distribución de la población 2 (una prueba de una cola).
3. Estadístico de prueba: $$z = \frac{T - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}$$
4. Región de rechazo:
   a. Para una prueba de dos colas, rechace H_0 si $z > z_{\alpha/2}$ o $z < -z_{\alpha/2}$.
   b. Para una prueba de una cola en la cola derecha, rechace H_0 si $z > z_\alpha$.
   c. Para una prueba de una cola en la cola izquierda, rechace H_0 si $z < -z_\alpha$.

O rechace H_0 si el valor p es $< \alpha$.

Los valores tabulados de z se encuentran en la tabla 3 del Apéndice.

EJEMPLO 11.4

Se realizó un experimento para comparar las resistencias de dos tipos de papel de estraza: uno es un papel de estraza estándar de un peso especificado y el otro es el mismo papel tratado con una sustancia química. Diez piezas de cada tipo de papel, seleccionadas al azar de la producción, produjeron las mediciones de resistencia que se muestran en la tabla 11.6. Pruebe la hipótesis nula de no diferencia en las distribuciones de resistencias para los dos tipos de papel contra la hipótesis alternativa de que el papel tratado tiende a ser más fuerte (es decir, su distribución de mediciones de resistencia se corre a la derecha de la distribución correspondiente para el papel no tratado).

TABLA 11.6 Mediciones de resistencia (y sus rangos) para dos tipos de papel

Estándar 1	Tratado 2
1.21 (2)	1.49 (15)
1.43 (12)	1.37 (7.5)
1.35 (6)	1.67 (20)
1.51 (17)	1.50 (16)
1.39 (9)	1.31 (5)
1.17 (1)	1.29 (3.5)
1.48 (14)	1.52 (18)
1.42 (11)	1.37 (7.5)
1.29 (3.5)	1.44 (13)
1.40 (10)	1.53 (19)
Suma de rango	$T_1 = 85.5$
	$T_1^* = n_1(n_1 + n_2 + 1) - T_1 = 210 - 85.5 = 124.5$

Solución Como los tamaños muestrales son iguales, estamos en libertad de decidir cuál de las dos muestras debe ser la muestra 1. Si se elige el tratamiento estándar como la primera muestra, se pueden clasificar 20 mediciones de elige y los valores de T_1 y T_1^* se ven en la parte inferior de la tabla. Como se desea detectar un corrimiento en las mediciones estándar (1) a la izquierda de las mediciones tratadas (2), se realiza una prueba de cola izquierda:

H_0: No hay diferencia en las distribuciones de resistencia
H_a: La distribución estándar se encuentra a la izquierda de la distribución tratada

y se usa $T = T_1$ como el estadístico de prueba, buscando un valor inusualmente pequeño de T.

Para hallar el valor crítico para una prueba de una cola con $\alpha = .05$, indicemos de la tabla 7a del Apéndice con $n_1 = n_2 = 10$. Usando la entrada de la tabla, se puede rechazar H_0 cuando

$T \leq 82$. Como el valor observado del estadístico de prueba es $T = 85.5$, no se puede rechazar H_0. Hay insuficiente evidencia para concluir que el papel de estraza tratado es más fuerte que el papel estándar.

Para usar la aproximación normal a la distribución de T, se calcula

$$\mu_T = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{10(21)}{2} = 105$$

y

$$\sigma_T^2 = \frac{n_1 n_2(n_1 + n_2 + 1)}{12} = \frac{10(10)(21)}{12} = 175$$

con $\sigma_T = \sqrt{175} = 13.23$. Entonces

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{85.5 - 105}{13.23} = -1.47$$

El valor p de una cola correspondiente a $z = -1.47$ es

$$\text{valor } p = P(z \leq -1.47) = .5 - .4292 = .0708$$

que es mayor que $\alpha = .05$. La conclusión es la misma. No se puede concluir que el papel de estraza tratado sea más fuerte que el papel estándar.

¿Cuándo se puede usar la prueba de la suma de rango de Wilcoxon en lugar de la prueba t no pareada de dos muestras? La prueba t de dos muestras funciona bien si los datos están normalmente distribuidos con varianzas iguales. Si hay duda respecto a estas suposiciones, se usa una gráfica de probabilidad normal para evaluar el grado de no normalidad y también una prueba F de dos muestras de varianzas muestrales para verificar la igualdad de varianzas. Si estos procedimientos indican ya sea no normalidad o desigualdad de varianza, entonces es apropiada la prueba de la suma de rango de Wilcoxon.

11.4 PRUEBA DE CORRIDAS CON UNA MUESTRA

Se puede usar una prueba de rango con signo, propuesta por Frank Wilcoxon para analizar el experimento de diferencia pareada de la sección 4.4 al considerar las diferencias pareadas de dos tratamientos, 1 y 2. Bajo la hipótesis nula de que no hay diferencias en las distribuciones para 1 y 2 se esperaría que (en promedio) la mitad de las diferencias en pares sean negativas y la mitad sean positivas; esto es, el número esperado de diferencias negativas entre pares sería $n/2$ (donde n es el número de pares). Además, se deduce que las diferencias positivas y negativas de igual magnitud absoluta deben presentarse con igual probabilidad. Si fuéramos a ordenar las diferencias de acuerdo con sus valores absolutos y de menor a mayor, las sumas de rango esperadas para las diferencias negativas y positivas sería igual. Las diferencias grandes en las sumas de los rangos, asignadas a las diferencias positivas y negativas, darían evidencia para indicar un cambio en lugar entre las distribuciones de respuestas para los dos tratamientos, 1 y 2.

Si la distribución 1 se corre a la derecha de la distribución 2, entonces se espera que más de las diferencias sean positivas y esto resulta en un número pequeño de diferencias negativas. Por tanto, para detectar esta alternativa de una cola, use la suma de rango T^- (la suma de los rangos de las diferencias negativas) y rechace la hipótesis nula para valores significativamente pequeños de T^-. Del mismo modo, si la distribución 1 se corre a la izquierda de la distribución 2, entonces se espera que más de las diferencias sean negativas y que el número de diferencias positivas sea pequeño. En consecuencia, para detectar esta alternativa de una cola, use T^+ (la suma de los rangos de las diferencias positivas) y rechace la hipótesis nula si T^+ es significativamente pequeña.

CÁLCULO DEL ESTADÍSTICO DE PRUEBA PARA LA PRUEBA DE RANGO CON SIGNO DE WILCOXON

1. Calcule las diferencias $(x_1 - x_2)$ para cada uno de los n pares. Las diferencias iguales a 0 se eliminan y el número de pares, n, se reduce de conformidad.
2. Ordene los **valores absolutos** de las diferencias asignando 1 a la más pequeña, 2 a la segunda más pequeña, y así sucesivamente. A las observaciones empatadas se les asigna el promedio de los rangos que se hubieran asignado sin empates.
3. Calcule la **suma de rango** para las diferencias **negativas** y marque este valor T^-. Del mismo modo, calcule T^+, la **suma de rango** para las diferencias **positivas**.

Para una **prueba de dos colas** use la **menor de estas dos cantidades T como un estadístico de prueba** para probar la hipótesis nula de que los dos histogramas de frecuencia relativa poblacional son idénticos. Cuanto menor sea el valor de T, mayor es el peso de evidencia a favor de rechazar la hipótesis nula. **Por tanto, se rechazará la hipótesis nula si T es menor o igual a algún valor, por ejemplo T_0.**

Para detectar la **alternativa de una cola**, esa **distribución 1 se corre a la derecha de la distribución 2, use la suma de rango T^-** de las diferencias negativas y rechace la hipótesis nula para valores pequeños de T^-, por ejemplo $T^+ \leq T_0$. Si desea detectar un **corrimiento de la distribución 2 a la derecha de la distribución 1, use la suma de rango T^+** de las diferencias positivas como estadístico de prueba y rechace la hipótesis nula para valores pequeños de T^+, por ejemplo, $T^+ \leq T_0$.

La probabilidad de que T sea menor o igual a algún valor T_0 se ha calculado para una combinación de tamaños muestrales y valores de T_0. Estas probabilidades, dadas en la tabla 8 del Apéndice, se pueden usar para hallar la región de rechazo para la prueba T.

Una versión abreviada de la tabla 8 se muestra en la tabla 11.7. En sentido horizontal en la parte superior de la tabla se ve el número de diferencias (el número de pares) n. Los valores de α para una prueba de una cola aparecen en la primera columna de la tabla. La segunda columna da valores de α para una prueba de dos colas. Las entradas de la tabla son los valores críticos de T. Usted recordará que el valor crítico de un estadístico de prueba es el valor que localiza la frontera de la región de rechazo.

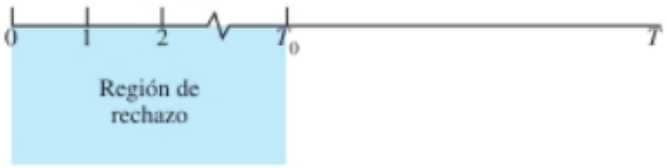
Por ejemplo, suponga que tenemos $n = 7$ pares y se realiza una prueba de dos colas de la hipótesis nula de que las dos distribuciones de frecuencia relativa poblacional son idénticas. Al comprobar la columna $n = 7$ de la tabla 11.7 y usar el segundo renglón (correspondiente a $\alpha = .05$ para una prueba de dos colas), se ve la entrada 2 (sombreada). Este valor es T_0, el valor crítico de T. Como ya se vio antes, cuanto menor sea el valor de T, mayor es la evidencia para rechazar la hipótesis nula. Por tanto, se rechazará la hipótesis nula para todos los valores de T menores o iguales a 2. La región de rechazo para la prueba de rango con signo de Wilcoxon para un experimento pareado es siempre de la forma: rechazar H_0 si $T \leq T_0$, donde T_0 es el valor crítico de T. La región de rechazo se muestra simbólicamente en la figura 11.2.

TABLA 11.7 **Versión abreviada de la tabla 8 del Apéndice; valores críticos de T**

Una cola	Dos colas	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$	$n = 11$
$\alpha = .050$	$\alpha = .10$	1	2	4	6	8	11	14
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8	11
$\alpha = .010$	$\alpha = .02$			0	2	3	5	7
$\alpha = .005$	$\alpha = .01$				0	2	3	5

Una cola	Dos colas	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$	$n = 17$
$\alpha = .050$	$\alpha = .10$	17	21	26	30	36	41
$\alpha = .025$	$\alpha = .05$	14	17	21	25	30	35
$\alpha = .010$	$\alpha = .02$	10	13	16	20	24	28
$\alpha = .005$	$\alpha = .01$	7	10	13	16	19	23

FIGURA 11.2
Región de rechazo para la prueba de rango con signo de Wilcoxon para un experimento pareado (rechazar H_0 si $T \leq T_0$)



PRUEBA DE RANGO CON SIGNO DE WILCOXON PARA UN EXPERIMENTO PAREADO

1. Hipótesis nula: H_0: Las dos distribuciones de frecuencia relativa poblacional son idénticas.
2. Hipótesis alternativa: H_a: Las dos distribuciones de frecuencia relativa poblacional difieren en ubicación (una prueba de dos colas); o H_a: La distribución de frecuencia relativa de la población 1 se corre a la derecha de la distribución de frecuencia relativa para la población 2 (una prueba de una cola).
3. Estadístico de prueba
   a. Para una prueba de dos colas, use T, la menor de la suma de rango para diferencias positivas y la suma de rango para diferencias negativas.
   b. Para una prueba de una cola (detectar la hipótesis alternativa descrita líneas antes), use la suma de rango T^- de las diferencias negativas.
4. Región de rechazo
   a. Para una prueba de dos colas, rechace H_0 si $T \leq T_0$, donde T_0 es el valor crítico dado en la tabla 8 del Apéndice.
   b. Para una prueba de una cola (detectar la hipótesis alternativa descrita líneas antes), use la suma de rango T^- de las diferencias negativas. Rechace H_0 si $T^- \leq T_0$.[†]

   [NOTA: se puede demostrar que $T^+ + T^- = \dfrac{n(n+1)}{2}$.]

EJEMPLO 11.5

Se realizó un experimento para comparar las densidades de pasteles elaborados de dos clases diferentes de harinas para pastel, A y B. Seis moldes para hornear pastel recibieron la masa A y seis la masa B. Esperando una variación en la temperatura del horno, el experimentador colocó un pastel A y uno B juntos en seis lugares diferentes del horno. Pruebe la hipótesis de que no hay diferencia en las distribuciones poblacionales de densidades de pastel para dos masas para pastel diferentes.

Solución Los datos (densidad en onzas por pulgada cúbica) y diferencias en densidad para seis pares de pasteles se dan en la tabla 11.8. La gráfica de caja de las diferencias de la figura 11.3 muestra un sesgo bastante fuerte y una diferencia muy grande en la cola derecha, lo cual indica que los datos pueden no satisfacer la suposición de normalidad. La muestra de dos diferencias es demasiado pequeña para tomar decisiones válidas acerca de normalidad y varianza constante. En esta situación, la prueba de rango con signo de Wilcoxon es la prueba más prudente a usar.

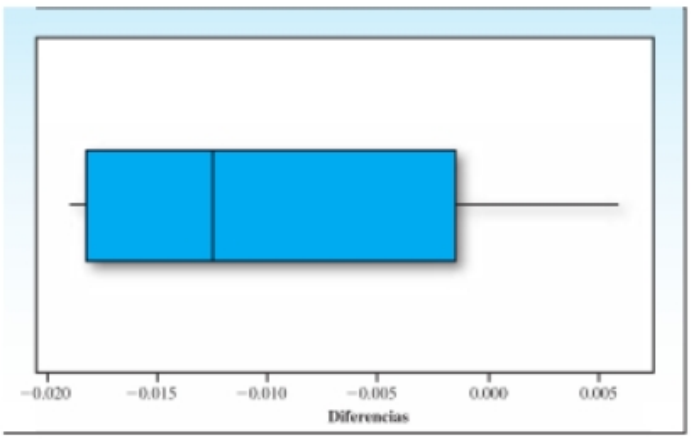
Al igual que con otras pruebas no paramétricas, la hipótesis nula a probar es que las dos distribuciones de frecuencia poblacionales de densidades de pastel son idénticas. La hipótesis alternativa, que implica una prueba de dos colas, es que las distribuciones son diferentes. Como la cantidad de datos es pequeña, se puede efectuar la prueba usando $\alpha = .10$. De la tabla 8 del Apéndice, el valor crítico de T para una prueba de dos colas, $\alpha = .10$, es $T_0 = 2$. Por tanto, se puede rechazar H_0 si $T \leq 2$.

[†] Para detectar un cambio de la distribución 2 a la derecha de la distribución 1, use la suma de rango T^+ de las diferencias positivas como estadístico de prueba y rechace H_0 si $T^+ \leq T_0$.

TABLA 11.8 **Densidades de seis pares de pasteles**

x_A	x_B	Diferencia $(x_A - x_B)$	Rango
.135	.129	.006	2
.102	.120	−.018	5
.098	.112	−.014	4
.141	.152	−.011	3
.131	.135	−.004	1
.144	.163	−.019	6

FIGURA 11.3
Gráfica de caja de diferencias para el ejemplo 11.5



Las diferencias $(x_1 - x_2)$ están calculadas y ordenadas de acuerdo con sus valores absolutos en la tabla 11.8. La suma de rangos positivos es $T^+ = 2$ y la suma de rangos negativos es $T^- = 19$. El estadístico de prueba es la más pequeña de estas dos sumas de rango, o $T = 2$. Como $T = 2$ cae en la región de rechazo, se puede rechazar H_0 y concluir que las dos distribuciones de frecuencia poblacional de densidades de pastel difieren.

Una salida impresa *MINITAB* de la prueba de rango con signo de Wilcoxon se da en la figura 11.4. En la sección 11.6 encontrará instrucciones para generar esta salida impresa. Observará que el valor del estadístico de prueba concuerda con los otros cálculos y el valor p indica que se puede rechazar H_0 al nivel de significancia de 10%.

FIGURA 11.4
Salida impresa *MINITAB* para el ejemplo 11.5

Prueba de rango con signo de Wilcoxon: diferencia

```
Test of median = 0.000000 versus median not = 0.000000

                N for   Wilcoxon           Estimated
           N    Test  Statistic        P     Median
Difference 6       6        2.0    0.093   -0.01100
```

Aproximación normal para la prueba de rango con signo de Wilcoxon

Aun cuando la tabla 8 del Apéndice tiene valores críticos para n de hasta 50, T^+, al igual que la prueba de rango con signo de Wilcoxon, estará distribuida en forma aproximadamente normal cuando la hipótesis nula sea verdadera y sea grande, por ejemplo 25 o más. Esto hace posible construir una prueba z de muestra grande, donde

$$E(T) = \frac{n(n+1)}{4}$$

$$\sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$$

Entonces el estadístico z

$$z = \frac{T^+ - E(T)}{\sigma_T} = \frac{T^+ - \dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}}$$

se usa como estadístico de prueba. Por tanto, para una prueba de dos colas y $\alpha = .05$, usted puede rechazar la hipótesis de distribuciones poblacionales idénticas cuando $|z| \geq 1.96$.

PRUEBA DE RANGO CON SIGNO DE WILCOXON DE MUESTRA GRANDE PARA UN EXPERIMENTO PAREADO: $n \geq 25$

1. Hipótesis nula: H_0: las distribuciones 1 y 2 de frecuencia relativa poblacional son idénticas.
2. Hipótesis alternativa: H_a: las dos distribuciones de frecuencia relativa poblacional difieren en ubicación (una prueba de dos colas); o bien, H_a: la distribución de frecuencia relativa poblacional 1 está corrida a la derecha (o izquierda) de la distribución de frecuencia relativa para la población 2 (una prueba de una cola).
3. Estadístico de prueba: $z = \dfrac{T^+ - [n(n+1)/4]}{\sqrt{[n(n+1)(2n+1)]/24}}$
4. Región de rechazo: rechazar H_0 si $z > z_{\alpha/2}$ o $z < -z_{\alpha/2}$ para una prueba de dos colas. Para una prueba de una cola, pónganse todas las α en una cola de la distribución z. Para detectar un cambio en la distribución 1 a la derecha de la distribución 2, rechazar H_0 cuando $z > z_\alpha$. Para detectar un cambio en la dirección opuesta, rechace H_0 si $z < -z_\alpha$.

Los valores tabulados de z se dan en la tabla 3 del Apéndice.

11.5 CORRELACIÓN POR RANGOS

En las secciones precedentes usamos rangos para indicar la magnitud relativa de observaciones en pruebas no paramétricas para la comparación de tratamientos. A continuación usaremos la misma técnica para probar una relación entre dos variables ordenadas. Dos coeficientes comunes de correlación de rango son r_s **de Spearman** y τ **de Kendall**. Presentaremos el r_s de Spearman porque su cálculo es idéntico al del coeficiente de correlación muestral r del capítulo 8.

Suponga que ocho profesores de ciencias de escuela básica han sido clasificados por un juez de acuerdo con su capacidad de enseñanza y todos han presentado un "examen nacional para maestros". Los datos se dan en la tabla 11.9. ¿Los datos sugieren un acuerdo entre la clasificación del juez y la calificación del examen? Esto es, ¿hay una correlación entre rangos y calificaciones de examen?

TABLA 11.9 **Clasificaciones y calificaciones de examen para ocho profesores**

Profesor	Clasificación del juez	Calificación de examen
1	7	44
2	4	72
3	2	69
4	6	70
5	1	93
6	3	82
7	8	67
8	5	80

Las dos variables de interés son clasificación del juez y calificación de examen. La primera ya está en forma de rango y las calificaciones se pueden ordenar de manera similar, como se ve en la tabla 11.10. Los rangos para las observaciones empatadas se obtienen al promediar los rangos que habrían tenido las observaciones empatadas si no se hubieran observado empates. El coeficiente de correlación de rango de Spearman, r_s, se calcula usando los rangos de las mediciones pareadas en las dos variables x y y en la fórmula para r (véase el capítulo 8).

TABLA 11.10 **Rangos de datos de la tabla 11.9**

Profesor	Clasificación del juez, x_i	Calificación de examen, y_i
1	7	1
2	4	5
3	2	3
4	6	4
5	1	8
6	3	7
7	8	2
8	5	6

COEFICIENTE DE CORRELACIÓN DE RANGO DE SPEARMAN

$$r_s = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

donde x_i y y_i representan los rangos del i-ésimo par de observaciones y

$$S_{xy} = \Sigma\,(x_i - \bar{x})(y_i - \bar{y}) = \Sigma\, x_i y_i - \frac{(\Sigma\, x_i)(\Sigma\, y_i)}{n}$$

$$S_{xx} = \Sigma\,(x_i - \bar{x})^2 = \Sigma\, x_i^2 - \frac{(\Sigma\, x_i)^2}{n}$$

$$S_{yy} = \Sigma\,(y_i - \bar{y})^2 = \Sigma\, y_i^2 - \frac{(\Sigma\, y_i)^2}{n}$$

Cuando no haya empates en las x observaciones o las y observaciones, la expresión para r_s algebraicamente se reduce a la expresión más sencilla

$$r_s = 1 - \frac{6\,\Sigma\, d_i^2}{n(n^2 - 1)} \qquad \text{donde } d_i = (x_i - y_i)$$

Si el número de empates es pequeño en comparación con el número de pares de datos, se produce poco error al usar esta fórmula breve.

EJEMPLO 11.6 Calcule r_s para los datos de la tabla 11.10.

Solución Las diferencias y cuadrados de diferencias entre los dos rangos se dan en la tabla 11.11. Sustituyendo valores en la fórmula para r_s, tenemos

$$r_s = 1 - \frac{6\,\Sigma\, d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(144)}{8(64 - 1)} = -.714$$

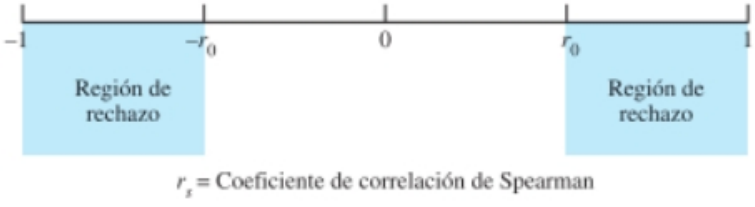
TABLA 11.11 **Diferencias y cuadrados de diferencias para los rangos del profesor**

Profesor	x_i	y_i	d_i	d_i^2
1	7	1	6	36
2	4	5	−1	1
3	2	3	−1	1
4	6	4	2	4
5	1	8	−7	49
6	3	7	−4	16
7	8	2	6	36
8	5	6	−1	1
Total				144

El coeficiente de correlación de rango de Spearman se usa como estadístico de prueba para verificar la hipótesis de que no hay asociación entre dos poblaciones. Se puede suponer que los n pares de observaciones (x_i, y_i) se han seleccionado al azar y, por tanto, ninguna asociación entre las poblaciones implica una asignación aleatoria de los n rangos dentro de cada muestra. Cada asignación aleatoria (para las dos muestras) representa un evento sencillo asociado con el experimento, y un valor de r_s se calcula para cada una. Entonces, es posible calcular la probabilidad de que r_s tome un valor absoluto grande debido sólo a una casualidad y, por tanto, sugiere una asociación entre poblaciones cuando no existe ninguna.

La región de rechazo para una prueba de dos colas se muestra en la figura 11.5. Si la hipótesis alternativa es que la correlación entre x y y sea negativa, se rechazaría H_0 para valores negativos de r_s cercanos a −1 (en la cola inferior de la figura 11.5). Del mismo modo, si la hipótesis alternativa es que la correlación entre x y y es positiva, se rechazaría H_0 para valores positivos grandes de r_s (en la cola superior de la figura 11.5).

FIGURA 11.5
Región de rechazo para una prueba de dos colas de la hipótesis nula de no asociación, usando la prueba de correlación de rango de Spearman



Los valores críticos de r_s se dan en la tabla 9 del Apéndice. Una versión abreviada se muestra en la tabla 11.12. En sentido horizontal en la tabla 11.12 (y la tabla 9 del Apéndice) están los valores registrados de α que podrían usarse para una prueba de una cola de la hipótesis nula de no asociación entre x y y. El número de pares de rango n aparece en el lado izquierdo de la tabla. Las entradas de la tabla dan el valor crítico r_0 para una prueba de una cola. Entonces, $P(r_s \geq r_0) = \alpha$.

Por ejemplo, suponga que tenemos $n = 8$ pares de rango y la hipótesis alternativa es que la correlación entre los rangos es positiva. Se desearía rechazar la hipótesis nula de no asociación sólo para los valores positivos grandes de r_s y se usaría una prueba de una cola. Si se consulta la tabla 11.12 y se usa el renglón correspondiente a $n = 8$ y la columna para $\alpha = .05$, se lee $r_0 = .643$. Por tanto, se puede rechazar H_0 para todos los valores de r_s mayores o iguales a .643.

La prueba se realiza exactamente en la misma forma si se desea probar sólo la hipótesis alternativa de que los rangos tienen correlación negativa. La única diferencia es que se rechazaría la hipótesis nula si $r_s \leq -.643$. Esto es, se usa el negativo del valor tabulado de r_0 para obtener el valor crítico de cola inferior.

TABLA 11.12 **Versión abreviada de la tabla 9 del Apéndice; para la prueba de correlación de rango de Spearman**

n	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	—
8	.643	.738	.833	.881
9	.600	.683	.783	.833
10	.564	.648	.745	.794
11	.523	.623	.736	.818
12	.497	.591	.703	.780
13	.475	.566	.673	.745
14	.457	.545		
15	.441	.525		
16	.425			
17	.412			
18	.399			
19	.388			
20	.377			

Para efectuar una prueba de dos colas se rechaza la hipótesis nula si $r_s \geq r_0$ o $r_s \leq -r_0$. El valor de α para la prueba es el doble del valor que se ve en la parte superior de la tabla. Por ejemplo, si $n = 8$ y se elige la columna .025, se rechaza H_0 si $r_s \geq .738$ o $r_s \leq -.738$. El valor α para la prueba es $2(.025) = .05$.

PRUEBA DE CORRELACIÓN DE RANGO DE SPEARMAN

1. Hipótesis nula: H_0: no hay asociación entre los pares de rangos.
2. Hipótesis alternativa: H_a: hay asociación entre los pares de rangos (una prueba de dos colas); o bien, H_a: la correlación entre los pares de rangos es positiva o negativa (una prueba de una cola).
3. Estadístico de prueba: $r_s = \dfrac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$

   donde x_i y y_i representan las filas del i-ésimo par de observaciones.
4. Región de rechazo: para una prueba de dos colas, rechace H_0 si $r_s \geq r_0$ o $r_s \leq -r_0$, donde r_0 se da en la tabla 9 del Apéndice. Duplique la probabilidad tabulada para obtener el valor de α para la prueba de dos colas. Para una prueba de una cola, rechace H_0 si $r_s \geq r_0$ (para una prueba de cola superior) o $r_s \leq -r_0$ (para una prueba de cola inferior). El valor α para una prueba de una cola es el valor que se muestra en la tabla 9 del Apéndice.

EJEMPLO 11.7 Pruebe la hipótesis de que no hay asociación entre las poblaciones para el ejemplo 11.6.

Solución El valor crítico de r_s para una prueba de una cola con $\alpha = .05$ y $n = 8$ es .643. Se puede suponer que una correlación entre la clasificación del juez y las calificaciones de examen de profesores no podrían ser posiblemente positivas. (Una clasificación baja significa buena enseñanza y debe estar asociada con una alta calificación de examen si el juez y la prueba miden la capacidad de enseñanza.) La hipótesis alternativa es que el **coeficiente de correlación de rango poblacional** ρ_s es menor que 0 y nos interesa una prueba estadística de una cola. Entonces, α para la prueba es el valor tabulado para .05 y se puede rechazar la hipótesis nula si $r_s \leq -.643$.

El valor calculado del estadístico de prueba, $r_s = -.714$, es menor que el valor crítico para $\alpha = .05$. En consecuencia, la hipótesis nula se rechaza al nivel de significancia $\alpha = .05$. Es

evidente que existe algún acuerdo entre las clasificaciones del juez y las calificaciones del examen. No obstante, debe observarse que este acuerdo podría existir cuando *ninguna* da una medida adecuada para medir la capacidad de enseñanza. Por ejemplo, la asociación podría existir si el juez y quienes formularon el examen de profesores tuvieran un concepto completamente erróneo, pero semejante, de las características de la buena enseñanza.

¿Qué es exactamente lo que mide r_s? El coeficiente de correlación de Spearman detecta no sólo una relación lineal entre dos variables, sino que también mide cualquier otra relación monotónica (y aumenta cuando x aumenta o y disminuye cuando x aumenta). Por ejemplo, si se calculó r_s para los dos conjuntos de datos de la tabla 11.13, ambos producirían un valor de $r_s = 1$ porque los rangos asignados para x y y en ambos casos concuerdan para todos los pares (x, y). Es importante recordar que un valor significativo de r_s indica una relación entre x y y que es creciente o decreciente, pero no necesariamente lineal.

TABLA 11.13 **Conjuntos de datos gemelos con $r_s = 1$**

x	$y = x^2$	x	$y = \log_{10}(x)$
1	1	10	1
2	4	100	2
3	9	1000	3
4	16	10,000	4
5	25	100,000	5
6	36	1,000,000	6

11.6 UTILIZACIÓN DE SOFTWARE DE ESTADÍSTICA PARA LA SOLUCIÓN DE CASOS

Procedimientos no paramétricos: *MINITAB*

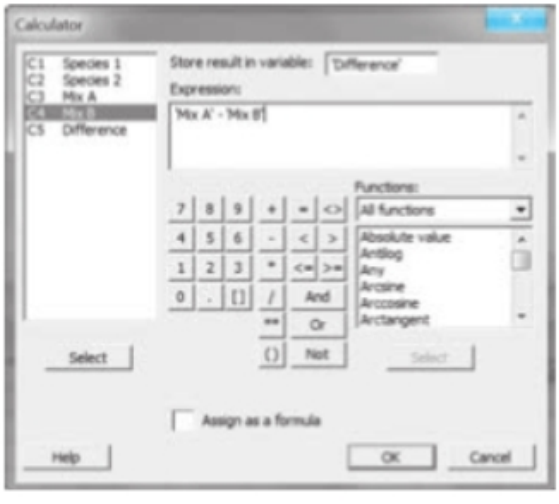
Aunque no hay opciones para procedimientos no paramétricos en *MS Excel*, muchos de estos procedimientos se encuentran disponibles en el paquete *MINITAB*, incluyendo la mayor parte de las pruebas estudiadas en este capítulo. Los cuadros de diálogo son conocidos para el usuario por ahora y veremos las pruebas en el orden presentado en el capítulo.

Para poner en práctica la prueba de la suma de rango de Wilcoxon para dos muestras aleatorias independientes introduzca los dos conjuntos de datos muestrales en dos columnas (por ejemplo, C1 y C2) de la hoja de trabajo *MINITAB*. El cuadro de diálogo de la figura 11.6 se genera usando **Stat ▶ Nonparametrics ▶ Mann-Whitney**. Seleccione C1 y C2 para la **First** y **Second Samples** e indique el coeficiente de confianza apropiado (para un intervalo de confianza) e hipótesis alternativa. Al hacer clic en **OK** se genera la salida de la figura 11.1.

La prueba del signo y la prueba de rango con signo de Wilcoxon para muestras pareadas se efectúan exactamente en la misma forma, con un cambio sólo en el último comando de la secuencia. Incluso los cuadros de diálogo son idénticos. Introduzca los datos en dos columnas de la hoja de trabajo *MINITAB* (usamos los datos de harina para pastel en la sección 11.4). Antes de que se pueda implementar cada prueba es preciso generar una columna de diferencias usando **Calc ▶ Calculator**, como se ve en la figura 11.7. Use **Stat ▶ Nonparametrics ▶ 1-Sample Sign** o **Stat ▶ Nonparametrics ▶ 1-Sample Wilcoxon** para generar el cuadro de diálogo apropiado que se muestra en la figura 11.8.

FIGURA 11.6

FIGURA 11.7



Recuerde que la mediana es el valor de una variable tal que 50% de los valores son más pequeños y 50% son más grandes. En consecuencia, si las dos distribuciones poblacionales son iguales, la mediana de las diferencias será 0. Esto es equivalente a la hipótesis nula

$$H_0: P(\text{diferencia positiva}) = P(\text{diferencia negativa}) = .5$$

FIGURA 11.8

empleada para la prueba del signo. Seleccione la columna de diferencias para el cuadro Variables y seleccione la prueba de la mediana igual a 0 con la alternativa apropiada. Haga clic en **OK** para obtener la salida impresa de cualquiera de las dos pruebas. La salida impresa de la ventana Session para la prueba del signo, ilustrada en la figura 11.9, indica una diferencia no significativa en las distribuciones de densidades para las dos harinas para pastel. Observe que el valor p (.2188) no es igual que el valor p para la prueba de rango con signo de Wilcoxon (.093 de la figura 11.4). No obstante, si se prueba al nivel de 5%, ambas pruebas producen diferencias no significativas.

FIGURA 11.9

Los procedimientos para implementar la prueba H de Kruskal-Wallis para k muestras independientes y la prueba F_r de Friedman para un diseño aleatorizado de bloques, son idénticos a los procedimientos empleados para sus equivalentes paramétricos. Repase los métodos descritos en la sección 6.4. Una vez que haya introducido los datos como se explica en esa sección, los comandos **Stat ▶ Nonparametrics ▶ Kruskal-Wallis** o **Stat ▶ Nonparametrics ▶ Friedman** van a generar un cuadro de diálogo en el que el usuario especifica la columna Response y la columna Factor o la columna de respuesta, la columna del tratamiento y la columna de bloque, respectivamente. Haga clic en **OK** para obtener las salidas impresas para estas pruebas no paramétricas.

Finalmente, se puede generar el coeficiente no paramétrico de correlación de rango r_s si se introducen los datos en dos columnas y se ordenan los datos usando **Data ▶ Rank**. Por ejemplo, los datos sobre la clasificación del juez y las calificaciones del examen se introdujeron en las columnas C6 y C7 de nuestra hoja de cálculo *MINITAB*. Como las clasificaciones del juez ya están ordenadas, sólo necesitamos ordenar C7 al seleccionar "Exam Score" y guardar los rangos en C8 [llamada "Rank (y)" en la figura 11.10]. Los comandos **Stat ▶ Basic Statistics ▶ Correlation** producirán ahora el coeficiente de correlación de rango cuando C6 y C8 se seleccionen. No obstante, el valor p que se ve en la *salida no produce* exactamente la misma prueba que los valores críticos de la tabla 11.12. Debe compararse el valor de r_s obtenido por el usuario con el valor tabulado para comprobar si hay una asociación significativa entre las dos variables.

FIGURA 11.10

EJERCICIOS

TÉCNICAS BÁSICAS

DATOS EX1516

11.1 Se realizó un experimento de diferencia pareada para comparar dos poblaciones. Los datos se muestran en la tabla siguiente. Use la prueba del signo para determinar si las distribuciones poblacionales son diferentes.

	Pares						
Población	1	2	3	4	5	6	7
1	8.9	8.1	9.3	7.7	10.4	8.3	7.4
2	8.8	7.4	9.0	7.8	9.9	8.1	6.9

a. Exprese las hipótesis nula y alternativa para la prueba.

b. Determine una región de rechazo apropiada con $\alpha \approx .01$.

c. Calcule el valor observado del estadístico de prueba.

d. ¿Los datos presentan suficiente evidencia para indicar que las poblaciones 1 y 2 son diferentes?

11.2 Observaciones de dos muestras aleatorias e independientes, tomadas de las poblaciones 1 y 2, se dan aquí. Use la prueba de la suma de rango de Wilcoxon para determinar si la población 1 está corrida a la izquierda de la población 2.

Muestra 1	1	3	2	3	5
Muestra 2	4	7	6	8	6

a. Exprese las hipótesis nula y alternativa a probar.

b. Ordene la muestra combinada de menor a mayor. Calcule T_1 y T_1^*.

c. ¿Cuál es la región de rechazo para $\alpha = .05$?

d. ¿Los datos dan suficiente evidencia para indicar que la población 1 está corrida a la izquierda de la población 2?

11.3 Suponga que desea detectar un corrimiento en la distribución 1 a la derecha de la distribución 2, con base en tamaños muestrales $n_1 = 12$ y $n_2 = 14$. Si $T_1 = 193$, ¿qué se concluye? Use $\alpha = .05$.

11.4 Suponga que se desea detectar una diferencia en las ubicaciones de dos distribuciones poblacionales, basadas en un experimento de diferencia pareada de $n = 30$ pares.

a. Dé las hipótesis nula y alternativa para la prueba de rango con signo de Wilcoxon.

b. Dé el estadístico de prueba.

c. Dé la región de rechazo para la prueba para $\alpha = .05$.

d. Si $T^+ = 249$, ¿cuáles son sus conclusiones? [NOTA: $T^+ + T^- = n(n+1)/2$.]

11.5 Se compararon tres tratamientos usando un diseño completamente aleatorizado. Los datos se muestran en la tabla siguiente.

DATOS EX1532

Tratamiento		
1	2	3
26	27	25
29	31	24
23	30	27
24	28	22
28	29	24
26	32	20
	30	21
	33	

¿Los datos dan suficiente evidencia para indicar una diferencia en ubicación para al menos dos de las distribuciones poblacionales? Pruebe usando el estadístico H de Kruskal-Wallis con $\alpha = .05$.

11.6 Se usa un diseño aleatorizado de bloques para comparar tres tratamientos en seis bloques.

DATOS EX1538

	Tratamiento		
Bloque	1	2	3
1	3.2	3.1	2.4
2	2.8	3.0	1.7
3	4.5	5.0	3.9
4	2.5	2.7	2.6
5	3.7	4.1	3.5
6	2.4	2.4	2.0

a. Use la prueba F_r de Friedman para detectar diferencias en ubicación entre las tres distribuciones de tratamiento. Pruebe usando $\alpha = .05$.

b. Encuentre el valor p aproximado para la prueba de la parte **a**.

c. Efectúe un análisis de varianza y dé la tabla ANOVA para el análisis.

d. Dé el valor del estadístico F para probar la igualdad de las tres medias de tratamiento.

e. Dé el valor p aproximado para el estadístico F de la parte **d**.

f. Compare los valores p para las pruebas de las partes **a** y **d**, y explique las implicaciones prácticas de la comparación.

11.7 Dé la región de rechazo en una prueba para detectar correlación de rango positiva si el número de pares de rangos es 16 y se tienen estos valores α:

a. $\alpha = .05$ **b.** $\alpha = .01$

11.8 Las siguientes observaciones pareadas se obtuvieron en dos variables x y y:

x	1.2	.8	2.1	3.5	2.7	1.5
y	1.0	1.3	.1	−.8	−.2	.6

a. Calcule el coeficiente de correlación de rango de Spearman r_s.

b. ¿Los datos presentan suficiente evidencia para indicar una correlación entre x y y? Pruebe usando $\alpha = .05$.

APLICACIONES

DATOS EX1520

11.9 Cocina de gourmet Dos gourmets, A y B, calificaron 22 comidas en una escala del 1 al 10. Los datos se muestran en la tabla. ¿Los datos dan suficiente evidencia para indicar que uno de los gourmets tiende a dar calificaciones más altas que el otro? Pruebe usando la prueba del signo con un valor de α cercano a .05.

Comida	A	B	Comida	A	B
1	6	8	12	8	5
2	4	5	13	4	2
3	7	4	14	3	3
4	8	7	15	6	8
5	2	3	16	9	10

Comida	A	B	Comida	A	B
6	7	4	17	9	8
7	9	9	18	4	6
8	7	8	19	4	3
9	2	5	20	5	4
10	4	3	21	3	2
11	6	9	22	5	3

a. Use las tablas binomiales del Apéndice para hallar la región de rechazo exacta para la prueba.

b. Use el estadístico z de muestra grande. (NOTA: Aun cuando la aproximación de muestra grande se sugiere para $n \geq 25$, funciona bastante bien para los valores de n tan pequeños como 15.)

c. Compare los resultados de las partes **a** y **b**.

11.10 Niveles de plomo en la sangre Un estudio publicado en la *American Journal of Public Health (Science News)*, primera publicación en seguir los niveles de plomo en la sangre de los aficionados al tiro al blanco con pistola en polígonos de tiro bajo techo y que cumplen con la ley, documenta un riesgo significativo de envenenamiento por plomo.[1] Se tomaron mediciones de exposición al plomo a 17 miembros del grupo de entrenamiento del cuerpo de policía antes, durante y después de un periodo de tres meses de instrucción de disparo en un polígono de tiro bajo techo y propiedad del Estado. Ningún recluta tenía niveles elevados de plomo en la sangre antes del entrenamiento, pero 15 de los 17 terminaron su entrenamiento con niveles de plomo considerados "elevados" por la Agencia para la Seguridad y la Salud en el Trabajo (OSHA). Si el uso de un polígono de tiro no causa un aumento en niveles de plomo en la sangre, entonces p, la probabilidad de que aumente el nivel de plomo en la sangre de una persona es menor o igual a .5. Pero si el uso del polígono de tiro bajo techo causa un aumento en los niveles de plomo en la sangre de una persona, entonces $p > .5$. Use la prueba del signo para determinar si el uso de un polígono de tiro bajo techo tiene el efecto de aumentar el nivel de plomo en la sangre con $\alpha = .05$. (SUGERENCIA: La aproximación normal a probabilidades binomiales es bastante precisa para $n = 17$.)

11.11 Enfermedad de Alzheimer En algunas pruebas de ancianos sanos un nuevo medicamento ha restaurado su memoria casi como la de los jóvenes. Pronto se probará en pacientes con enfermedad de Alzheimer, esa fatal enfermedad del cerebro que destruye la mente. Según el doctor Gary Lynch, de la Universidad de California en Irvine, el medicamento, llamado ampakina CX-516, acelera las señales entre las células cerebrales que parecen agudizar significativamente la memoria.[2] En una prueba preliminar en estudiantes de poco más de 20 años y en hombres de entre 65 y 70 años de edad, los resultados fueron particularmente sorprendentes. Después de recibir dosis moderadas de este medicamento, las personas de entre 65 y 70 años de edad calificaron casi tan alto como los jóvenes. Los datos siguientes son los números de sílabas sin sentido recordadas después de 5 minutos, para 10 hombres de poco más de 20 años de edad y 10 señores de entre 65 y 70 años. Use la prueba de la suma de rango de Wilcoxon para determinar si las distribuciones para el número de sílabas sin sentido recordadas son iguales para estos dos grupos.

20	3	6	4	8	7	1	1	2	7	8
65–70	1	0	4	1	2	5	0	2	2	3

DATOS EX1510

11.12 Comparación de mariscales de campo de la NFL ¿Cómo se compara Aaron Rodgers, mariscal de campo de los ganadores del Súper Tazón de 2011, los Vikingos de Minnesota, con Drew Brees, mariscal de campo de los ganadores del Súper Tazón de 2010, los Santos de Nueva Orleans? La tabla que sigue muestra el número de pases completos para cada atleta durante la temporada de futbol de 2010 de la NFL:[3]

Aaron Rodgers			Drew Brees		
19	21	7	27	37	25
19	15	25	28	34	29
34	27	19	30	27	35
12	22		33	29	22
27	26		24	23	
18	21		21	24	

Use la prueba de suma de rangos de Wilcoxon para analizar los datos y probar para ver si las distribuciones poblacionales para el número de pases completos difieren para los dos mariscales de campo. Use $\alpha = .05$.

DATOS EX1559

11.13 Brillantez de papel Los valores codificados para una medida de la brillantez del papel (reflectividad de la luz), preparados por dos procesos diferentes, aparecen en la tabla para muestras de nueve observaciones tomadas al azar de cada uno de los dos procesos. ¿Los datos presentan suficiente evidencia para indicar una diferencia en las mediciones de la brillantez para los dos procesos? Use una prueba paramétrica y una no paramétrica y compare sus resultados.

Proceso	Brillantez								
A	6.1	9.2	8.7	8.9	7.6	7.1	9.5	8.3	9.0
B	9.1	8.2	8.6	6.9	7.5	7.9	8.3	7.8	8.9

DATOS EX1565

11.14 Matemáticas y arte La tabla muestra las calificaciones de un grupo de 15 estudiantes en matemáticas y arte. Use la prueba de rango con signo de Wilcoxon para determinar si las calificaciones medianas para estos estudiantes difieren significativamente para las dos materias.

Estudiante	Matemáticas	Arte	Estudiante	Matemáticas	Arte
1	22	53	9	62	55
2	37	68	10	65	74
3	36	42	11	66	68
4	38	49	12	56	64
5	42	51	13	66	67
6	58	65	14	67	73
7	58	51	15	62	65
8	60	71			

11.15 Investigación del SIDA Unos científicos han demostrado que una vacuna recién inventada puede proteger a monos Rhesus contra infecciones causadas por un virus estrechamente relacionado con el virus de inmunodeficiencia humana (VIH) causante del SIDA. En su trabajo, Ronald C. Resrosiers y sus colegas del Centro Regional de Investigación de Primates de Nueva Inglaterra, aplicaron a cada uno de $n = 6$ monos Rhesus cinco inoculaciones con la vacuna del virus de inmunodeficiencia de simios (VIS). Una semana después de la última vacuna, cada mono recibió una inyección de VIS vivo. Dos de los seis monos vacunados no mostraron evidencia de infección por el VIS hasta por año y medio después de la inyección con el VIS.[4] Los científicos pudieron aislar el virus del VIS de los otros cuatro monos vacunados, aun cuando estos animales no mostraron signos de la enfermedad. ¿Esta información contiene suficiente evidencia para indicar que la vacuna es eficaz para proteger al menos contra el VIS? Use $\alpha = .10$.

DATOS EX1578

11.16 Fatiga de trabajadores Para investigar métodos para reducir la fatiga entre los empleados cuyos trabajos comprenden un monótono procedimiento de ensamblaje, a 12 empleados seleccionados al azar se les pidió que realizaran su trabajo normal en una de tres condiciones de prueba. Como medida de fatiga el experimentador usó el número de paradas de la línea de ensamblaje durante un periodo de cuatro horas para cada condición de prueba.

	Condiciones		
Empleado	1	2	3
1	31	22	26
2	20	15	23
3	26	21	18
4	31	22	32
5	12	16	18
6	22	29	34
7	28	17	26
8	15	9	12
9	41	31	46
10	19	19	25
11	31	34	41
12	18	11	21

a. ¿Qué tipo de diseño experimental se ha empleado en este experimento?

b. Use la prueba no paramétrica apropiada para determinar si la distribución de paradas de la línea de ensamble (y en consecuencia la fatiga del trabajador) difiere para estas tres condiciones. Pruebe al nivel de significancia de 5%.

11.17 Clasificación de mariscales de campo Se hizo una clasificación de mariscales de campo, de los ocho mejores equipos de la National Football League, al encuestar a varios periodistas de deportes y entrenadores profesionales. Esta "clasificación verdadera" se muestra a continuación, junto con "mi clasificación".

	Mariscal de campo							
	A	B	C	D	E	F	G	H
Clasificación verdadera	1	2	3	4	5	6	7	8
Mi clasificación	3	1	4	5	2	8	6	7

a. Calcule r_s.

b. ¿Los datos indican una correlación positiva entre mi clasificación y la de los expertos? Pruebe al nivel de significancia de 5%.

Fuentes

1. *Science News* 136 (Agosto de 1989):126.
2. "Alzheimer's Test Set for New Memory Drug," *The Press-Enterprise* (Riverside, CA), 18 de noviembre de 1997, p. A-4.
3. "Aaron Rodgers #12 QB," http://sports.espn.go.com/nfl/players/gamelog?playerId=8439 y "Drew Brees #9 QB," http://sports.espn.go.com/nfl/players/gamelog?playerId=2580, 18 de marzo de 2011.
4. *Science News*, 1989, p. 116.

Apéndice
Tablas

CONTENIDO

Tabla 1 Probabilidades binomiales acumulativas

Tabla 2 Probabilidades acumulativas de Poisson

Tabla 3 Áreas bajo la curva normal

Tabla 4 Valores críticos de t

Tabla 5 Valores críticos de chi cuadrada

Tabla 6 Puntos porcentuales de la distribución F

Tabla 7 Valores críticos de T para la prueba de suma de rango de Wilcoxon, $n_1 \leq n_2$

Tabla 8 Valores críticos de T para la prueba de rango con signo de Wilcoxon, $n = 5(1)50$

Tabla 9 Valores críticos del coeficiente de correlación de rango de Spearman para una prueba de una cola

Tabla 10 Números aleatorios

Tabla 11 Puntos porcentuales del rango de Student, $q_\alpha(k, df)$

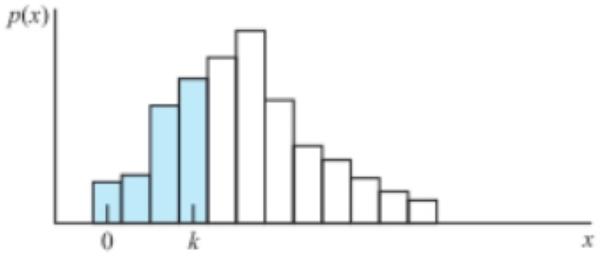


TABLA 1 Probabilidades binomiales acumulativas
Los valores tabulados son $P(x \leq k) = p(0) + p(1) + \cdots + p(k)$.
(Los cálculos están redondeados al tercer lugar decimal.)

$n = 2$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.980	.902	.810	.640	.490	.360	.250	.160	.090	.040	.010	.002	.000	0
1	1.000	.998	.990	.960	.910	.840	.750	.640	.510	.360	.190	.098	.020	1
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	2

$n = 3$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.970	.857	.729	.512	.343	.216	.125	.064	.027	.008	.001	.000	.000	0
1	1.000	.993	.972	.896	.784	.648	.500	.352	.216	.104	.028	.007	.000	1
2	1.000	1.000	.999	.992	.973	.936	.875	.784	.657	.488	.271	.143	.030	2
3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	3

$n = 4$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.961	.815	.656	.410	.240	.130	.062	.026	.008	.002	.000	.000	.000	0
1	.999	.986	.948	.819	.652	.475	.312	.179	.084	.027	.004	.000	.000	1
2	1.000	1.000	.996	.973	.916	.821	.688	.525	.348	.181	.052	.014	.001	2
3	1.000	1.000	1.000	.998	.992	.974	.938	.870	.760	.590	.344	.185	.039	3
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	4

TABLA 1 *(continuación)*

$n = 5$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000	0
1	.999	.977	.919	.737	.528	.337	.188	.087	.031	.007	.000	.000	.000	1
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000	2
3	1.000	1.000	1.000	.993	.969	.913	.812	.663	.472	.263	.081	.023	.001	3
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049	4
5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	5

$n = 6$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.941	.735	.531	.262	.118	.047	.016	.004	.001	.000	.000	.000	.000	0
1	.999	.967	.886	.655	.420	.233	.109	.041	.011	.002	.000	.000	.000	1
2	1.000	.998	.984	.901	.744	.544	.344	.179	.070	.017	.001	.000	.000	2
3	1.000	1.000	.999	.983	.930	.821	.656	.456	.256	.099	.016	.002	.000	3
4	1.000	1.000	1.000	.998	.989	.959	.891	.767	.580	.345	.114	.033	.001	4
5	1.000	1.000	1.000	1.000	.999	.996	.984	.953	.882	.738	.469	.265	.059	5
6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	6

$n = 7$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.932	.698	.478	.210	.082	.028	.008	.002	.000	.000	.000	.000	.000	0
1	.998	.956	.850	.577	.329	.159	.062	.019	.004	.000	.000	.000	.000	1
2	1.000	.996	.974	.852	.647	.420	.227	.096	.029	.005	.000	.000	.000	2
3	1.000	1.000	.997	.967	.874	.710	.500	.290	.126	.033	.003	.000	.000	3
4	1.000	1.000	1.000	.995	.971	.904	.773	.580	.353	.148	.026	.004	.000	4
5	1.000	1.000	1.000	1.000	.996	.981	.938	.841	.671	.423	.150	.044	.002	5
6	1.000	1.000	1.000	1.000	1.000	.998	.992	.972	.918	.790	.522	.302	.068	6
7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	7

$n = 8$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.923	.663	.430	.168	.058	.017	.004	.001	.000	.000	.000	.000	.000	0
1	.997	.943	.813	.503	.255	.106	.035	.009	.001	.000	.000	.000	.000	1
2	1.000	.994	.962	.797	.552	.315	.145	.050	.011	.001	.000	.000	.000	2
3	1.000	1.000	.995	.944	.806	.594	.363	.174	.058	.010	.000	.000	.000	3
4	1.000	1.000	1.000	.990	.942	.826	.637	.406	.194	.056	.005	.000	.000	4
5	1.000	1.000	1.000	.999	.989	.950	.855	.685	.448	.203	.038	.006	.000	5
6	1.000	1.000	1.000	1.000	.999	.991	.965	.894	.745	.497	.187	.057	.003	6
7	1.000	1.000	1.000	1.000	1.000	.999	.996	.983	.942	.832	.570	.337	.077	7
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	8

TABLA 1 *(continuación)*

$n = 9$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.914	.630	.387	.134	.040	.010	.002	.000	.000	.000	.000	.000	.000	0
1	.997	.929	.775	.436	.196	.071	.020	.004	.000	.000	.000	.000	.000	1
2	1.000	.992	.947	.738	.463	.232	.090	.025	.004	.000	.000	.000	.000	2
3	1.000	.999	.992	.914	.730	.483	.254	.099	.025	.003	.000	.000	.000	3
4	1.000	1.000	.999	.980	.901	.733	.500	.267	.099	.020	.001	.000	.000	4
5	1.000	1.000	1.000	.997	.975	.901	.746	.517	.270	.086	.008	.001	.000	5
6	1.000	1.000	1.000	1.000	.996	.975	.910	.768	.537	.262	.053	.008	.000	6
7	1.000	1.000	1.000	1.000	1.000	.996	.980	.929	.804	.564	.225	.071	.003	7
8	1.000	1.000	1.000	1.000	1.000	1.000	.998	.990	.960	.866	.613	.370	.086	8
9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	9

$n = 10$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.904	.599	.349	.107	.028	.006	.001	.000	.000	.000	.000	.000	.000	0
1	.996	.914	.736	.376	.149	.046	.011	.002	.000	.000	.000	.000	.000	1
2	1.000	.988	.930	.678	.383	.167	.055	.012	.002	.000	.000	.000	.000	2
3	1.000	.999	.987	.879	.650	.382	.172	.055	.011	.001	.000	.000	.000	3
4	1.000	1.000	.998	.967	.850	.633	.377	.166	.047	.006	.000	.000	.000	4
5	1.000	1.000	1.000	.994	.953	.834	.623	.367	.150	.033	.002	.000	.000	5
6	1.000	1.000	1.000	.999	.989	.945	.828	.618	.350	.121	.013	.001	.000	6
7	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.322	.070	.012	.000	7
8	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.624	.264	.086	.004	8
9	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.893	.651	.401	.096	9
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	10

$n = 11$

	p												
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.895	.569	.314	.086	.020	.004	.000	.000	.000	.000	.000	.000	.000	0
1	.995	.898	.697	.322	.113	.030	.006	.001	.000	.000	.000	.000	.000	1
2	1.000	.985	.910	.617	.313	.119	.033	.006	.001	.000	.000	.000	.000	2
3	1.000	.998	.981	.839	.570	.296	.113	.029	.004	.000	.000	.000	.000	3
4	1.000	1.000	.997	.950	.790	.533	.274	.099	.022	.002	.000	.000	.000	4
5	1.000	1.000	1.000	.988	.922	.754	.500	.246	.078	.012	.000	.000	.000	5
6	1.000	1.000	1.000	.998	.978	.901	.726	.467	.210	.050	.003	.000	.000	6
7	1.000	1.000	1.000	1.000	.996	.971	.887	.704	.430	.161	.019	.002	.000	7
8	1.000	1.000	1.000	1.000	.999	.994	.967	.881	.687	.383	.090	.015	.000	8
9	1.000	1.000	1.000	1.000	1.000	.999	.994	.970	.887	.678	.303	.102	.005	9
10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.980	.914	.686	.431	.105	10
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	11

TABLA 1 *(continuación)*

$n = 12$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.886	.540	.282	.069	.014	.002	.000	.000	.000	.000	.000	.000	.000	0
1	.994	.882	.659	.275	.085	.020	.003	.000	.000	.000	.000	.000	.000	1
2	1.000	.980	.889	.558	.253	.083	.019	.003	.000	.000	.000	.000	.000	2
3	1.000	.998	.974	.795	.493	.225	.073	.015	.002	.000	.000	.000	.000	3
4	1.000	1.000	.996	.927	.724	.438	.194	.057	.009	.001	.000	.000	.000	4
5	1.000	1.000	.999	.981	.882	.665	.387	.158	.039	.004	.000	.000	.000	5
6	1.000	1.000	1.000	.996	.961	.842	.613	.335	.118	.019	.001	.000	.000	6
7	1.000	1.000	1.000	.999	.991	.943	.806	.562	.276	.073	.004	.000	.000	7
8	1.000	1.000	1.000	1.000	.998	.985	.927	.775	.507	.205	.026	.002	.000	8
9	1.000	1.000	1.000	1.000	1.000	.997	.981	.917	.747	.442	.111	.020	.000	9
10	1.000	1.000	1.000	1.000	1.000	1.000	.997	.980	.915	.725	.341	.118	.006	10
11	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.986	.931	.718	.460	.114	11
12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	12

$n = 15$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.860	.463	.206	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.990	.829	.549	.167	.035	.005	.000	.000	.000	.000	.000	.000	.000	1
2	1.000	.964	.816	.398	.127	.027	.004	.000	.000	.000	.000	.000	.000	2
3	1.000	.995	.944	.648	.297	.091	.018	.002	.000	.000	.000	.000	.000	3
4	1.000	.999	.987	.836	.515	.217	.059	.009	.001	.000	.000	.000	.000	4
5	1.000	1.000	.998	.939	.722	.403	.151	.034	.004	.000	.000	.000	.000	5
6	1.000	1.000	1.000	.982	.869	.610	.304	.095	.015	.001	.000	.000	.000	6
7	1.000	1.000	1.000	.996	.950	.787	.500	.213	.050	.004	.000	.000	.000	7
8	1.000	1.000	1.000	.999	.985	.905	.696	.390	.131	.018	.000	.000	.000	8
9	1.000	1.000	1.000	1.000	.996	.966	.849	.597	.278	.061	.002	.000	.000	9
10	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.164	.013	.001	.000	10
11	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.352	.056	.005	.000	11
12	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.602	.184	.036	.000	12
13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.833	.451	.171	.010	13
14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.794	.537	.140	14
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	15

TABLA 1 *(continuación)*

$n = 20$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.818	.358	.122	.012	.001	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.983	.736	.392	.069	.008	.001	.000	.000	.000	.000	.000	.000	.000	1
2	.999	.925	.677	.206	.035	.004	.000	.000	.000	.000	.000	.000	.000	2
3	1.000	.984	.867	.411	.107	.016	.001	.000	.000	.000	.000	.000	.000	3
4	1.000	.997	.957	.630	.238	.051	.006	.000	.000	.000	.000	.000	.000	4
5	1.000	1.000	.989	.804	.416	.126	.021	.002	.000	.000	.000	.000	.000	5
6	1.000	1.000	.998	.913	.608	.250	.058	.006	.000	.000	.000	.000	.000	6
7	1.000	1.000	1.000	.968	.772	.416	.132	.021	.001	.000	.000	.000	.000	7
8	1.000	1.000	1.000	.990	.887	.596	.252	.057	.005	.000	.000	.000	.000	8
9	1.000	1.000	1.000	.997	.952	.755	.412	.128	.017	.001	.000	.000	.000	9
10	1.000	1.000	1.000	.999	.983	.872	.588	.245	.048	.003	.000	.000	.000	10
11	1.000	1.000	1.000	1.000	.995	.943	.748	.404	.113	.010	.000	.000	.000	11
12	1.000	1.000	1.000	1.000	.999	.979	.868	.584	.228	.032	.000	.000	.000	12
13	1.000	1.000	1.000	1.000	1.000	.994	.942	.750	.392	.087	.002	.000	.000	13
14	1.000	1.000	1.000	1.000	1.000	.998	.979	.874	.584	.196	.011	.000	.000	14
15	1.000	1.000	1.000	1.000	1.000	1.000	.994	.949	.762	.370	.043	.003	.000	15
16	1.000	1.000	1.000	1.000	1.000	1.000	.999	.984	.893	.589	.133	.016	.000	16
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.965	.794	.323	.075	.001	17
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.992	.931	.608	.264	.017	18
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.988	.878	.642	.182	19
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	20

TABLA 1 *(continuación)*

$n = 25$

	p													
k	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	k
0	.778	.277	.072	.004	.000	.000	.000	.000	.000	.000	.000	.000	.000	0
1	.974	.642	.271	.027	.002	.000	.000	.000	.000	.000	.000	.000	.000	1
2	.998	.873	.537	.098	.009	.000	.000	.000	.000	.000	.000	.000	.000	2
3	1.000	.966	.764	.234	.033	.002	.000	.000	.000	.000	.000	.000	.000	3
4	1.000	.993	.902	.421	.090	.009	.000	.000	.000	.000	.000	.000	.000	4
5	1.000	.999	.967	.617	.193	.029	.002	.000	.000	.000	.000	.000	.000	5
6	1.000	1.000	.991	.780	.341	.074	.007	.000	.000	.000	.000	.000	.000	6
7	1.000	1.000	.998	.891	.512	.154	.022	.001	.000	.000	.000	.000	.000	7
8	1.000	1.000	1.000	.953	.677	.274	.054	.004	.000	.000	.000	.000	.000	8
9	1.000	1.000	1.000	.983	.811	.425	.115	.013	.000	.000	.000	.000	.000	9
10	1.000	1.000	1.000	.994	.902	.586	.212	.034	.002	.000	.000	.000	.000	10
11	1.000	1.000	1.000	.998	.956	.732	.345	.078	.006	.000	.000	.000	.000	11
12	1.000	1.000	1.000	1.000	.983	.846	.500	.154	.017	.000	.000	.000	.000	12
13	1.000	1.000	1.000	1.000	.994	.922	.655	.268	.044	.002	.000	.000	.000	13
14	1.000	1.000	1.000	1.000	.998	.966	.788	.414	.098	.006	.000	.000	.000	14
15	1.000	1.000	1.000	1.000	1.000	.987	.885	.575	.189	.017	.000	.000	.000	15
16	1.000	1.000	1.000	1.000	1.000	.996	.946	.726	.323	.047	.000	.000	.000	16
17	1.000	1.000	1.000	1.000	1.000	.999	.978	.846	.488	.109	.002	.000	.000	17
18	1.000	1.000	1.000	1.000	1.000	1.000	.993	.926	.659	.220	.009	.000	.000	18
19	1.000	1.000	1.000	1.000	1.000	1.000	.998	.971	.807	.383	.033	.001	.000	19
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.910	.579	.098	.007	.000	20
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.967	.766	.236	.034	.000	21
22	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.991	.902	.463	.127	.002	22
23	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.973	.729	.358	.026	23
24	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.928	.723	.222	24
25	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	25

TABLA 2 Probabilidades acumulativas de Poisson

Los valores tabulados son $P(x \leq k) = p(0) + p(1) + \cdots + p(k)$.
(Los cálculos están redondeados al tercer lugar decimal.)

	μ										
k	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.5
0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368	.223
1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736	.558
2	1.000	.999	.996	.992	.986	.977	.966	.953	.937	.920	.809
3		1.000	1.000	.999	.998	.997	.994	.991	.987	.981	.934
4				1.000	1.000	1.000	.999	.999	.998	.996	.981
5							1.000	1.000	1.000	.999	.996
6										1.000	.999
7											1.000

	μ										
k	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0
0	.135	.082	.050	.033	.018	.011	.007	.004	.003	.002	.001
1	.406	.287	.199	.136	.092	.061	.040	.027	.017	.011	.007
2	.677	.544	.423	.321	.238	.174	.125	.088	.062	.043	.030
3	.857	.758	.647	.537	.433	.342	.265	.202	.151	.112	.082
4	.947	.891	.815	.725	.629	.532	.440	.358	.285	.224	.173
5	.983	.958	.916	.858	.785	.703	.616	.529	.446	.369	.301
6	.995	.986	.966	.935	.889	.831	.762	.686	.606	.563	.450
7	.999	.996	.988	.973	.949	.913	.867	.809	.744	.673	.599
8	1.000	.999	.996	.990	.979	.960	.932	.894	.847	.792	.729
9		1.000	.999	.997	.992	.983	.968	.946	.916	.877	.830
10			1.000	.999	.997	.993	.986	.975	.957	.933	.901
11				1.000	.999	.998	.995	.989	.980	.966	.947
12					1.000	.999	.998	.996	.991	.984	.973
13						1.000	.999	.998	.996	.993	.987
14							1.000	.999	.999	.997	.994
15								1.000	.999	.999	.998
16									1.000	1.000	.999
17											1.000

TABLA 2 *(continuación)*

	μ								
k	7.5	8.0	8.5	9.0	9.5	10.0	12.0	15.0	20.0
0	.001	.000	.000	.000	.000	.000	.000	.000	.000
1	.005	.003	.002	.001	.001	.000	.000	.000	.000
2	.020	.014	.009	.006	.004	.003	.001	.000	.000
3	.059	.042	.030	.021	.015	.010	.002	.000	.000
4	.132	.100	.074	.055	.040	.029	.008	.001	.000
5	.241	.191	.150	.116	.089	.067	.020	.003	.000
6	.378	.313	.256	.207	.165	.130	.046	.008	.000
7	.525	.453	.386	.324	.269	.220	.090	.018	.001
8	.662	.593	.523	.456	.392	.333	.155	.037	.002
9	.776	.717	.653	.587	.522	.458	.242	.070	.005
10	.862	.816	.763	.706	.645	.583	.347	.118	.011
11	.921	.888	.849	.803	.752	.697	.462	.185	.021
12	.957	.936	.909	.876	.836	.792	.576	.268	.039
13	.978	.966	.949	.926	.898	.864	.682	.363	.066
14	.990	.983	.973	.959	.940	.917	.772	.466	.105
15	.995	.992	.986	.978	.967	.951	.844	.568	.157
16	.998	.996	.993	.989	.982	.973	.899	.664	.221
17	.999	.998	.997	.995	.991	.986	.937	.749	.297
18	1.000	.999	.999	.998	.996	.993	.963	.819	.381
19		1.000	.999	.999	.998	.997	.979	.875	.470
20			1.000	1.000	.999	.998	.988	.917	.559
21					1.000	.999	.994	.947	.644
22						1.000	.997	.967	.721
23							.999	.981	.787
24							.999	.989	.843
25							1.000	.994	.888
26								.997	.922
27								.998	.948
28								.999	.966
29								1.000	.978
30									.987
31									.992
32									.995
33									.997
34									.999
35									.999
36									1.000

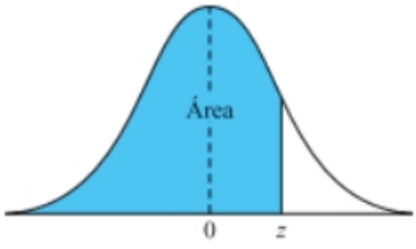


TABLA 3 Áreas bajo la curva normal

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

TABLA 3 *(continuación)*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5 871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

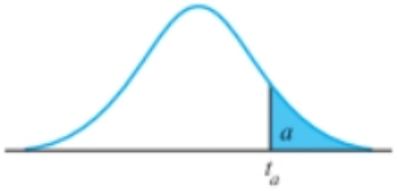


TABLA 4
Valores críticos de *t*

gl	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	*gl*
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
∞	1.282	1.645	1.960	2.326	2.576	∞

FUENTE: De "Table of Percentage Points of the t-Distribution", *Biometrika,* 32 (1941): 300. Reproducida con permiso de los fideicomisarios de *Biometrika.*

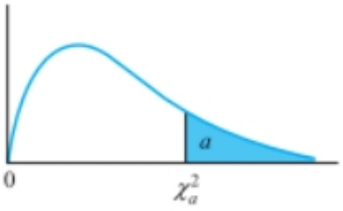


TABLA 5
Valores críticos de chi cuadrada

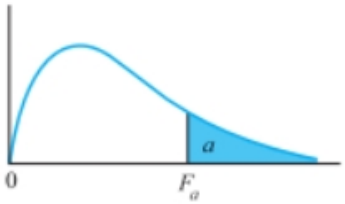
gl	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$
1	.0000393	.0001571	.0009821	.0039321	.0157908
2	.0100251	.0201007	.0506356	.102587	.210720
3	.0717212	.114832	.215795	.351846	.584375
4	.206990	.297110	.484419	.710721	1.063623
5	.411740	.554300	.831211	1.145476	1.61031
6	.675727	.872085	1.237347	1.63539	2.20413
7	.989265	1.239043	1.68987	2.16735	2.83311
8	1.344419	1.646482	2.17973	2.73264	3.48954
9	1.734926	2.087912	2.70039	3.32511	4.16816
10	2.15585	2.55821	3.24697	3.94030	4.86518
11	2.60321	3.05347	3.81575	4.57481	5.57779
12	3.07382	3.57056	4.40379	5.22603	6.30380
13	3.56503	4.10691	5.00874	5.89186	7.04150
14	4.07468	4.66043	5.62872	6.57063	7.78953
15	4.60094	5.22935	6.26214	7.26094	8.54675
16	5.14224	5.81221	6.90766	7.96164	9.31223
17	5.69724	6.40776	7.56418	8.67176	10.0852
18	6.26481	7.01491	8.23075	9.39046	10.8649
19	6.84398	7.63273	8.90655	10.1170	11.6509
20	7.43386	8.26040	9.59083	10.8508	12.4426
21	8.03366	8.89720	10.28293	11.5913	13.2396
22	8.64272	9.54249	10.9823	12.3380	14.0415
23	9.26042	10.19567	11.6885	13.0905	14.8479
24	9.88623	10.8564	12.4011	13.8484	15.6587
25	10.5197	11.5240	13.1197	14.6114	16.4734
26	11.1603	12.1981	13.8439	15.3791	17.2919
27	11.8076	12.8786	14.5733	16.1513	18.1138
28	12.4613	13.5648	15.3079	16.9279	18.9392
29	13.1211	14.2565	16.0471	17.7083	19.7677
30	13.7867	14.9535	16.7908	18.4926	20.5992
40	20.7065	22.1643	24.4331	26.5093	29.0505
50	27.9907	29.7067	32.3574	34.7642	37.6886
60	35.5346	37.4848	40.4817	43.1879	46.4589
70	43.2752	45.4418	48.7576	51.7393	55.3290
80	51.1720	53.5400	57.1532	60.3915	64.2778
90	59.1963	61.7541	65.6466	69.1260	73.2912
100	67.3276	70.0648	74.2219	77.9295	82.3581

FUENTE: De "Tables of the Percentage Points of the χ^2 Distribution", *Biometrika Tables for Statisticians*, vol. 1, 3a. ed. (1966). Reproducida con permiso de los fideicomisarios de *Biometrika*.

TABLA 5
(continuación)

$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$	*gl*
2.70554	3.84146	5.02389	6.63490	7.87944	1
4.60517	5.99147	7.37776	9.21034	10.5966	2
6.25139	7.81473	9.34840	11.3449	12.8381	3
7.77944	9.48773	11.1433	13.2767	14.8602	4
9.23635	11.0705	12.8325	15.0863	16.7496	5
10.6446	12.5916	14.4494	16.8119	18.5476	6
12.0170	14.0671	16.0128	18.4753	20.2777	7
13.3616	15.5073	17.5346	20.0902	21.9550	8
14.6837	16.9190	19.0228	21.6660	23.5893	9
15.9871	18.3070	20.4831	23.2093	25.1882	10
17.2750	19.6751	21.9200	24.7250	26.7569	11
18.5494	21.0261	23.3367	26.2170	28.2995	12
19.8119	22.3621	24.7356	27.6883	29.8194	13
21.0642	23.6848	26.1190	29.1413	31.3193	14
22.3072	24.9958	27.4884	30.5779	32.8013	15
23.5418	26.2962	28.8485	31.9999	34.2672	16
24.7690	27.8571	30.1910	33.4087	35.7185	17
25.9894	28.8693	31.5264	34.8053	37.1564	18
27.2036	30.1435	32.8523	36.1908	38.5822	19
28.4120	31.4104	34.1696	37.5662	39.9968	20
29.6151	32.6705	35.4789	38.9321	41.4010	21
30.8133	33.9244	36.7807	40.2894	42.7956	22
32.0069	35.1725	38.0757	41.6384	44.1813	23
33.1963	36.4151	39.3641	42.9798	45.5585	24
34.3816	37.6525	40.6465	44.3141	46.9278	25
35.5631	38.8852	41.9232	45.6417	48.2899	26
36.7412	40.1133	43.1944	46.9630	49.6449	27
37.9159	41.3372	44.4607	48.2782	50.9933	28
39.0875	42.5569	45.7222	49.5879	52.3356	29
40.2560	43.7729	46.9792	50.8922	53.6720	30
51.8050	55.7585	59.3417	63.6907	66.7659	40
63.1671	67.5048	71.4202	76.1539	79.4900	50
74.3970	79.0819	83.2976	88.3794	91.9517	60
85.5271	90.5312	95.0231	100.425	104.215	70
96.5782	101.879	106.629	112.329	116.321	80
107.565	113.145	118.136	124.116	128.299	90
118.498	124.342	129.561	135.807	140.169	100

TABLA 6 Puntos porcentuales de la distribución *F*



		df_1								
gl_2	*a*	1	2	3	4	5	6	7	8	9
1	.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	.050	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
	.025	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
	.010	4052	4999.5	5403	5625	5764	5859	5928	5982	6022
	.005	16211	20000	21615	22500	23056	23437	23715	23925	24091
2	.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
	.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	.005	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4
3	.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	.010	34.12	30.82	29.46	28.71	28.24	27.91	27.64	27.49	27.35
	.005	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.005	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.005	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77
6	.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	.005	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39
7	.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	.005	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
8	.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	.005	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34
9	.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	.005	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54

FUENTE: Parte de "Tables of Percentage Points of the Inverted Beta (F) Distribution", *Biometrika*, vol. 33 (1943); por M. Merrington y C. M. Thompson y de la tabla 18 de *Biometrika Tables for Statisticians*, vol. 1, Cambridge University Press, 1954, editado por E. S. Pearson y H. O. Hartley. Reproducida con permiso de los autores, editores y los fideicomisarios de *Biometrika*.

TABLA 6 *(continuación)*

gl_1											
10	12	15	20	24	30	40	60	120	∞	a	gl_2
60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33	.100	1
241.9	243.9	245.9	248.0	249.1	250.1	251.2	252.2	253.3	254.3	.050	
968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018	.025	
6056	6106	6157	6209	6235	6261	6287	6313	6339	6366	.010	
24224	24426	24630	24836	24940	25044	25148	25253	25359	25465	.005	
9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49	.100	2
19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50	.050	
39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50	.025	
99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50	.010	
199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5	.005	
5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13	.100	3
8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53	.050	
14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90	.025	
27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13	.010	
43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83	.005	
3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76	.100	4
5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63	.050	
8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26	.025	
14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46	.010	
20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32	.005	
3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10	.100	5
4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36	.050	
6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02	.025	
10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02	.010	
13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14	.005	
2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72	.100	6
4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67	.050	
5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85	.025	
7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88	.010	
10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88	.005	
2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47	.100	7
3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23	.050	
4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14	.025	
6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65	.010	
8.38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08	.005	
2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29	.100	8
3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93	.050	
4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67	.025	
5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86	.010	
7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95	.005	
2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16	.100	9
3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71	.050	
3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33	.025	
5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31	.010	
6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	5.19	.005	

TABLA 6 *(continuación)*

		gl_1								
gl_2	a	1	2	3	4	5	6	7	8	9
10	.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	.005	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
11	.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	.005	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54
12	.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	.005	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
13	.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	.005	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94
14	.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	.005	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72
15	.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	.005	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
16	.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	.005	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38
17	.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	.005	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25
18	.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	.005	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14
19	.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	.005	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04
20	.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	.005	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96

TABLA 6 *(continuación)*

gl_1											
10	12	15	20	24	30	40	60	120	∞	a	gl_2
2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06	.100	10
2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54	.050	
3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08	.025	
4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91	.010	
5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.86	4.75	4.64	.005	
2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97	.100	11
2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40	.050	
3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88	.025	
4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60	.010	
5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23	.005	
2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90	.100	12
2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30	.050	
3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72	.025	
4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36	.010	
5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90	.005	
2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85	.100	13
2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21	.050	
3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60	.025	
4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17	.010	
4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.76	3.65	.005	
2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80	.100	14
2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13	.050	
3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49	.025	
3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00	.010	
4.60	4.43	4.25	4.06	3.96	3.86	3.76	3.66	3.55	3.44	.005	
2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76	.100	15
2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07	.050	
3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40	.025	
3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87	.010	
4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26	.005	
2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72	.100	16
2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01	.050	
2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32	.025	
3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75	.010	
4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11	.005	
2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69	.100	17
2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96	.050	
2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25	.025	
3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65	.010	
4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98	.005	
1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66	.100	18
2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92	.050	
2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19	.025	
3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57	.010	
4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87	.005	
1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63	.100	19
2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88	.050	
2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13	.025	
3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49	.010	
3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78	.005	
1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61	.100	20
2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84	.050	
2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09	.025	
3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42	.010	
3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69	.005	

TABLA 6 *(continuación)*

		gl_1								
gl_2	a	1	2	3	4	5	6	7	8	9
21	.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	.005	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88
22	.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	.005	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81
23	.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	.005	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75
24	.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	.005	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
25	.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	.005	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
26	.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	.005	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60
27	.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	.005	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.005	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.005	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.005	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45

TABLA 6 *(continuación)*

gl_1											
10	12	15	20	24	30	40	60	120	∞	a	gl_2
1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59	.100	21
2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81	.050	
2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04	.025	
3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36	.010	
3.77	3.60	3.43	3.24	3.15	3.05	2.95	2.84	2.73	2.61	.005	
1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57	.100	22
2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78	.050	
2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00	.025	
3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31	.010	
3.70	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55	.005	
1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55	.100	23
2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76	.050	
2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97	.025	
3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26	.010	
3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48	.005	
1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53	.100	24
2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73	.050	
2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94	.025	
3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21	.010	
3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43	.005	
1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52	.100	25
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71	.050	
2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91	.025	
3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17	.010	
3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38	.005	
1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50	.100	26
2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69	.050	
2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88	.025	
3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13	.010	
3.49	3.33	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33	.005	
1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49	.100	27
2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67	.050	
2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85	.025	
3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10	.010	
3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.41	2.29	.005	
1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48	.100	28
2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65	.050	
2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83	.025	
3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06	.010	
3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.25	.005	
1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47	.100	29
2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64	.050	
2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81	.025	
3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03	.010	
3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.21	.005	
1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46	.100	30
2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62	.050	
2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79	.025	
2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01	.010	
3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18	.005	

TABLA 6 *(continuación)*

		gl_1								
gl_2	a	1	2	3	4	5	6	7	8	9
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.005	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.005	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
120	.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	.050	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	.005	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
∞	.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	.050	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.63
	.025	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	.005	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62

TABLA 6 *(continuación)*

gl_1											
10	12	15	20	24	30	40	60	120	∞	a	gl_2
1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38	.100	40
2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51	.050	
2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64	.025	
2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80	.010	
3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93	.005	
1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29	.100	60
1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39	.050	
2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48	.025	
2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60	.010	
2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.83	1.69	.005	
1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19	.100	120
1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25	.050	
2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31	.025	
2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38	.010	
2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.43	.005	
1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00	.100	∞
1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00	.050	
2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00	.025	
2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00	.010	
2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00	.005	

TABLA 7 Valores críticos de T para la prueba de suma de rango de Wilcoxon, $n_1 \leq n_2$

TABLA 7a
Valores críticos de cola izquierda a 5%

	n_1													
n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	—	6												
4	—	6	11											
5	3	7	12	19										
6	3	8	13	20	28									
7	3	8	14	21	29	39								
8	4	9	15	23	31	41	51							
9	4	10	16	24	33	43	54	66						
10	4	10	17	26	35	45	56	69	82					
11	4	11	18	27	37	47	59	72	86	100				
12	5	11	19	28	38	49	62	75	89	104	120			
13	5	12	20	30	40	52	64	78	92	108	125	142		
14	6	13	21	31	42	54	67	81	96	112	129	147	166	
15	6	13	22	33	44	56	69	84	99	116	133	152	171	192

TABLA 7b
Valores críticos de cola izquierda a 2.5%

	n_1													
n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4	—	—	10											
5	—	6	11	17										
6	—	7	12	18	26									
7	—	7	13	20	27	36								
8	3	8	14	21	29	38	49							
9	3	8	14	22	31	40	51	62						
10	3	9	15	23	32	42	53	65	78					
11	3	9	16	24	34	44	55	68	81	96				
12	4	10	17	26	35	46	58	71	84	99	115			
13	4	10	18	27	37	48	60	73	88	103	119	136		
14	4	11	19	28	38	50	62	76	91	106	123	141	160	
15	4	11	20	29	40	52	65	79	94	110	127	145	164	184

Fuente: Datos de "An Extended Table of Critical Values for the Mann-Whitney (Wilcoxon) Two-Sample Statistic" por Roy C. Milton, pp. 925-934 en *Journal of the American Statistical Association,* vol. 59, núm. 307, septiembre de 1964. Reimpresa con permiso de *Journal of the American Statistical Association.* Copyright 1964 por la American Statistical Association. Todos los derechos reservados.

TABLA 7c
Valores críticos de cola izquierda a 1%

	n_1													
n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	—	—												
4	—	—	—											
5	—	—	10	16										
6	—	—	11	17	24									
7	—	6	11	18	25	34								
8	—	6	12	19	27	35	45							
9	—	7	13	20	28	37	47	59						
10	—	7	13	21	29	39	49	61	74					
11	—	7	14	22	30	40	51	63	77	91				
12	—	8	15	23	32	42	53	66	79	94	109			
13	3	8	15	24	33	44	56	68	82	97	113	130		
14	3	8	16	25	34	45	58	71	85	100	116	134	152	
15	3	9	17	26	36	47	60	73	88	103	120	138	156	176

TABLA 7d
Valores críticos de cola izquierda a .5%

	n_1												
n_2	3	4	5	6	7	8	9	10	11	12	13	14	15
3	—												
4	—	—											
5	—	—	15										
6	—	10	16	23									
7	—	10	16	24	32								
8	—	11	17	25	34	42							
9	6	11	18	26	35	45	56						
10	6	12	19	27	37	47	58	71					
11	6	12	20	28	38	49	61	73	87				
12	7	13	21	30	40	51	63	76	90	105			
13	7	13	22	31	41	53	65	79	93	109	125		
14	7	14	22	32	43	54	67	81	96	112	129	147	
15	8	15	23	33	44	56	69	84	99	115	133	151	171

TABLA 8
Valores críticos de *T* para la prueba de rango con signo de Wilcoxon, *n* = 5(1)50

Una cola	Dos colas	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$\alpha = .050$	$\alpha = .10$	1	2	4	6	8	11
$\alpha = .025$	$\alpha = .05$		1	2	4	6	8
$\alpha = .010$	$\alpha = .02$			0	2	3	5
$\alpha = .005$	$\alpha = .01$				0	2	3

Una cola	Dos colas	$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$\alpha = .050$	$\alpha = .10$	14	17	21	26	30	36
$\alpha = .025$	$\alpha = .05$	11	14	17	21	25	30
$\alpha = .010$	$\alpha = .02$	7	10	13	16	20	24
$\alpha = .005$	$\alpha = .01$	5	7	10	13	16	19

Una cola	Dos colas	$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$\alpha = .050$	$\alpha = .10$	41	47	54	60	68	75
$\alpha = .025$	$\alpha = .05$	35	40	46	52	59	66
$\alpha = .010$	$\alpha = .02$	28	33	38	43	49	56
$\alpha = .005$	$\alpha = .01$	23	28	32	37	43	49

Una cola	Dos colas	$n = 23$	$n = 24$	$n = 25$	$n = 26$	$n = 27$	$n = 28$
$\alpha = .050$	$\alpha = .10$	83	92	101	110	120	130
$\alpha = .025$	$\alpha = .05$	73	81	90	98	107	117
$\alpha = .010$	$\alpha = .02$	62	69	77	85	93	102
$\alpha = .005$	$\alpha = .01$	55	68	68	76	84	92

Una cola	Dos colas	$n = 29$	$n = 30$	$n = 31$	$n = 32$	$n = 33$	$n = 34$
$\alpha = .050$	$\alpha = .10$	141	152	163	175	188	201
$\alpha = .025$	$\alpha = .05$	127	137	148	159	171	183
$\alpha = .010$	$\alpha = .02$	111	120	130	141	151	162
$\alpha = .005$	$\alpha = .01$	100	109	118	128	138	149

Una cola	Dos colas	$n = 35$	$n = 36$	$n = 37$	$n = 38$	$n = 39$
$\alpha = .050$	$\alpha = .10$	214	228	242	256	271
$\alpha = .025$	$\alpha = .05$	195	208	222	235	250
$\alpha = .010$	$\alpha = .02$	174	186	198	211	224
$\alpha = .005$	$\alpha = .01$	160	171	183	195	208

Una cola	Dos colas	$n = 40$	$n = 41$	$n = 42$	$n = 43$	$n = 44$	$n = 45$
$\alpha = .050$	$\alpha = .10$	287	303	319	336	353	371
$\alpha = .025$	$\alpha = .05$	264	279	295	311	327	344
$\alpha = .010$	$\alpha = .02$	238	252	267	281	297	313
$\alpha = .005$	$\alpha = .01$	221	234	248	262	277	292

Una cola	Dos colas	$n = 46$	$n = 47$	$n = 48$	$n = 49$	$n = 50$
$\alpha = .050$	$\alpha = .10$	389	408	427	446	466
$\alpha = .025$	$\alpha = .05$	361	379	397	415	434
$\alpha = .010$	$\alpha = .02$	329	345	362	380	398
$\alpha = .005$	$\alpha = .01$	307	323	339	356	373

FUENTE: De "Some Rapid Approximate Statistical Procedures" (1964) 28, por F. Wilcoxon y R. A. Wilcox. Reproducida con el bondadoso permiso de Lederle Laboratories, división de American Cyanamid Company.

TABLA 9
Valores críticos del coeficiente de correlación de rango de Spearman para una prueba de una cola

n	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$
5	.900	—	—	—
6	.829	.886	.943	—
7	.714	.786	.893	—
8	.643	.738	.833	.881
9	.600	.683	.783	.833
10	.564	.648	.745	.794
11	.523	.623	.736	.818
12	.497	.591	.703	.780
13	.475	.566	.673	.745
14	.457	.545	.646	.716
15	.441	.525	.623	.689
16	.425	.507	.601	.666
17	.412	.490	.582	.645
18	.399	.476	.564	.625
19	.388	.462	.549	.608
20	.377	.450	.534	.591
21	.368	.438	.521	.576
22	.359	.428	.508	.562
23	.351	.418	.496	.549
24	.343	.409	.485	.537
25	.336	.400	.475	.526
26	.329	.392	.465	.515
27	.323	.385	.456	.505
28	.317	.377	.448	.496
29	.311	.370	.440	.487
30	.305	.364	.432	.478

FUENTE: De "Distribution of Sums of Squares of Rank Differences for Small Samples" por E. G. Olds, *Annals of Mathematical Statistics* 9 (1938). Reproducida con el permiso del editor, *Annals of Mathematical Statistics.*

TABLA 10 Números aleatorios

	Columna													
Fila	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
3	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
8	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
10	84575	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
18	01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
21	48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
24	29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	36188	18510
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
26	81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	79666	95725
27	29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
28	00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760	64585	96096	98253
29	05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470	66520	34693	90449
30	91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618
31	00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808	76655	62028	76630
32	00725	69884	62797	56170	86324	88072	76222	36086	84637	93161	76038	65855	77919	88006
33	69011	65795	95876	55293	18988	27354	26575	08625	40801	59920	29841	80150	12777	48501
34	25976	57948	29888	88604	67917	48708	18912	82271	65424	69774	33611	54262	85963	03547
35	09763	83473	73577	12908	30883	18317	28290	35797	05998	41688	34952	37888	38917	88050
36	91567	42595	27958	30134	04024	86385	29880	99730	55536	84855	29080	09250	79656	73211
37	17955	56349	90999	49127	20044	59931	06115	20542	18059	02008	73708	83517	36103	42791
38	46503	18584	18845	49618	02304	51038	20655	58727	28168	15475	56942	53389	20562	87338
39	92157	89634	94824	78171	84610	82834	09922	25417	44137	48413	25555	21246	35509	20468
40	14577	62765	35605	81263	39667	47358	56873	56307	61607	49518	89656	20103	77490	18062
41	98427	07523	33362	64270	01638	92477	66969	98420	04880	45585	46565	04102	46880	45709
42	34914	63976	88720	82765	34476	17032	87589	40836	32427	70002	70663	88863	77775	69348
43	70060	28277	39475	46473	23219	53416	94970	25832	69975	94884	19661	72828	00102	66794
44	53976	54914	06990	67245	68350	82948	11398	42878	80287	88267	47363	46634	06541	97809
45	76072	29515	40980	07391	58745	25774	22987	80059	39911	96189	41151	14222	60697	59583
46	90725	52210	83974	29992	65831	38857	50490	83765	55657	14361	31720	57375	56228	41546
47	64364	67412	33339	31926	14883	24413	59744	92351	97473	89286	35931	04110	23726	51900
48	08962	00358	31662	25388	61642	34072	81249	35648	56891	69352	48373	45578	78547	81788
49	95012	68379	93526	70765	10592	04542	76463	54328	02349	17247	28865	14777	62730	92277
50	15664	10493	20492	38391	91132	21999	59516	81652	27195	48223	46751	22923	32261	85653

FUENTE: De *Handbook of Tables for Probabiliy and Statistics,* 2a. ed., por William H. Beyer (CRC Press). Usada con permiso de William H. Beyer.

TABLA 10 *(continuación)*

	Columna													
Fila	1	2	3	4	5	6	7	8	9	10	11	12	13	14
51	16408	81899	04153	53381	79401	21438	83035	92350	36693	31238	59649	91754	72772	02338
52	18629	81953	05520	91962	04739	13092	97662	24822	94730	06496	35090	04822	86774	98289
53	73115	35101	47498	87637	99016	71060	88824	71013	18735	20286	23153	72924	35165	43040
54	57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812	12515	98931	91202
55	30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68668	30429	70735	25499
56	16631	35006	85900	98275	32388	52390	16815	69298	82732	38480	73817	32523	41961	44437
57	96773	20206	42559	78985	05300	22164	24369	54224	35033	19687	11052	91491	60383	19746
58	38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318	29686	03387	59846
59	31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325
60	78919	19474	23632	27889	47914	02584	37680	20801	72152	39339	34806	08930	85001	87820
61	03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833	25570	38818	46920
62	74426	33278	43972	10119	89917	15665	52872	73823	73144	88662	88970	74492	51805	99378
63	09066	00903	20795	95452	92648	45454	09552	88815	16553	51125	79375	97596	16296	66092
64	42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689	05974	52468	16834
65	16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510	70625	28725	34191
66	21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240	15957	16572	06004
67	21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995	43805	33386	21597
68	55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525	42786	05269	92532
69	44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911	25650	12682	73572
70	91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203	71795	99533	50501
71	91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867	14951	91696	85065
72	50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405	56087	82790	70925
73	65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202	94617	23772	07896
74	27504	96131	83944	41575	10573	08619	64482	73923	36152	05184	94142	25299	84387	34925
75	37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330	74301	00275	48280
76	11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84081	66938	93654	59894
77	37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651	50245	34971	52924
78	46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345	81073	49106	79860
79	30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339	58861	74818	46942
80	63798	64995	46583	09785	44160	78128	83991	42865	92520	83531	80377	35909	81250	54238
81	82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870	52689	51275	83556
82	21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194	52799	28225	85762
83	60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535	12133	14645	23541
84	43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434	98227	21824	19585
85	97656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534	03862	78095	50136
86	03299	01221	05418	38982	55758	92237	26759	86367	21216	98442	08303	56613	91511	75928
87	79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076	72811	22717	50585
88	85636	68335	47539	03129	65651	11977	02510	26113	99447	68645	34327	15152	55230	93448
89	18039	14367	61337	06177	12143	46609	32989	74014	64708	00533	35398	58408	13261	47908
90	08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639	82163	60859	75567
91	79556	29068	04142	16268	15387	12856	66227	38358	22478	73373	88732	09443	82558	05250
92	92608	82674	27072	32534	17075	27698	98204	63863	11951	34648	88022	56148	34925	57031
93	23982	25835	40055	67006	12293	02753	14827	23235	35071	99704	37543	11601	35503	85171
94	09915	96306	05908	97901	28395	14186	00821	80703	70426	75647	76310	88717	37890	40129
95	59037	33300	26695	62247	69927	76123	50842	43834	86654	70959	79725	93872	28117	19233
96	42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	80799	76536	71255	64239
97	46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	53203	18098	47625	88684
98	03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	06216	95787	42579	90730
99	86591	81482	52667	61582	14972	90053	89534	76036	49199	43716	97548	04379	46370	28672
100	38534	01715	94964	87288	65680	43772	39560	12918	86737	62738	19636	51132	25739	56947

TABLA 11a
Puntos porcentuales del rango de Student, $q_{.05}(k, gl)$; puntos de 5% superior

	k									
gl	2	3	4	5	6	7	8	9	10	11
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36
15	3.01	3.67	4.08	4.37	4.60	4.78	4.94	5.08	5.20	5.31
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26
17	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73
120	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64
∞	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55

TABLA 11a
(continuación)

k									
12	13	14	15	16	17	18	19	20	*gl*
51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83	59.56	1
14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57	16.77	2
9.95	10.15	10.35	10.52	10.69	10.84	10.98	11.11	11.24	3
8.21	8.37	8.52	8.66	8.79	8.91	9.03	9.13	9.23	4
7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	5
6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59	6
6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17	7
6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	8
5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64	9
5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47	10
5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33	11
5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	12
5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11	13
5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03	14
5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96	15
5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	16
5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84	17
5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	18
5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75	19
5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71	20
5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59	24
5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47	30
4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36	40
4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	60
4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	120
4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01	∞

FUENTE: De *Biometrika Tables for Statisticians,* vol. 1, 3a. ed., editado por E. S. Pearson y H. O. Hartley (Cambridge University Press, 1966). Reproducida con permiso de los fideicomisarios de *Biometrika.*

TABLA 11b
Puntos porcentuales del rango de Student, $q_{.01}(k, gl)$; puntos de 1% superior

	k									
gl	2	3	4	5	6	7	8	9	10	11
1	90.03	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6	253.2
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69	32.59
3	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69	17.13
4	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27	12.57
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55
8	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85
40	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69
60	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23

TABLA 11b
(continuación)

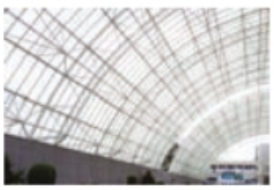
k									
12	13	14	15	16	17	18	19	20	*gl*
260.0	266.2	271.8	277.0	281.8	286.3	290.0	294.3	298.0	1
33.40	34.13	34.81	35.43	36.00	36.53	37.03	37.50	37.95	2
17.53	17.89	18.22	18.52	18.81	19.07	19.32	19.55	19.77	3
12.84	13.09	13.32	13.53	13.73	13.91	14.08	14.24	14.40	4
10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93	5
9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	6
8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	7
8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	8
7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57	9
7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23	10
7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	11
7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	12
6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55	13
6.77	6.87	6.96	7.05	7.13	7.20	7.27	7.33	7.39	14
6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	15
6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	16
6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05	17
6.41	6.50	6.58	6.65	6.72	6.79	6.85	6.91	6.97	18
6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	19
6.28	6.37	6.45	6.52	6.59	6.65	6.71	6.77	6.82	20
6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	24
5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	30
5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21	40
5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01	60
5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83	120
5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65	∞

FUENTE: De *Biometrika Tables for Statisticians*, vol. 1, 3a. ed., editado por E. S. Pearson y H. O. Hartley (Cambridge University Press, 1966). Reproducida con permiso de los fideicomisarios de *Biometrika*.

50 AÑOS UNITEC

Campus Atizapán

Campus Cuitláhuac



Campus Ecatepec

Campus Guadalajara

Campus León

Campus Marina

Campus Sur

Campus Toluca

ISBN-13: 978-607-526-560-5
ISBN-10: 607-526-560-0

CENGAGE Learning®

Visita nuestro sitio en http://latinoamerica.cengage.com